

Tesis de Posgrado

Mecanismo de unión a ADN del dominio C-terminal de la proteína E2 del papilomavirus humano, principal regulador de la expresión génica viral

Ferreiro, Diego U.

2003

Tesis presentada para obtener el grado de Doctor en Ciencias Químicas de la Universidad de Buenos Aires

Este documento forma parte de la colección de tesis doctorales y de maestría de la Biblioteca Central Dr. Luis Federico Leloir, disponible en digital.bl.fcen.uba.ar. Su utilización debe ser acompañada por la cita bibliográfica con reconocimiento de la fuente.

This document is part of the doctoral theses collection of the Central Library Dr. Luis Federico Leloir, available in digital.bl.fcen.uba.ar. It should be used accompanied by the corresponding citation acknowledging the source.

Cita tipo APA:

Ferreiro, Diego U.. (2003). Mecanismo de unión a ADN del dominio C-terminal de la proteína E2 del papilomavirus humano, principal regulador de la expresión génica viral. Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires.

http://digital.bl.fcen.uba.ar/Download/Tesis/Tesis_3638_Ferreiro.pdf

Cita tipo Chicago:

Ferreiro, Diego U.. "Mecanismo de unión a ADN del dominio C-terminal de la proteína E2 del papilomavirus humano, principal regulador de la expresión génica viral". Tesis de Doctor. Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires. 2003.

http://digital.bl.fcen.uba.ar/Download/Tesis/Tesis_3638_Ferreiro.pdf

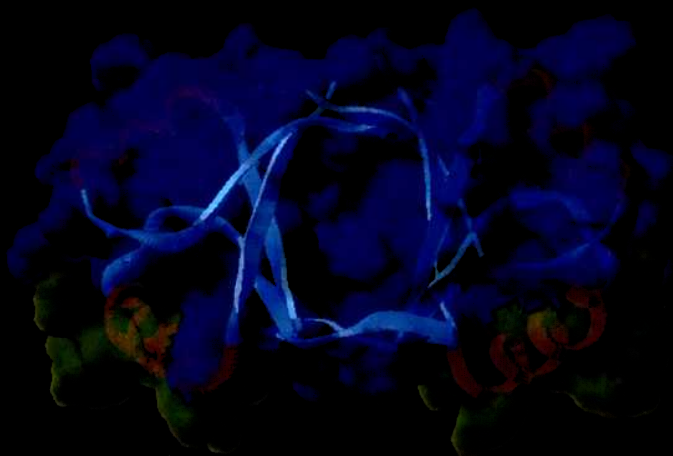
EXACTAS UBA

Facultad de Ciencias Exactas y Naturales



UBA

Universidad de Buenos Aires



AUTOR:
LIC. DIEGO U. FERREIRO

DIRECTOR:
DR. GONZALO PRAT GAY

**TESIS PRESENTADA PARA
OPTAR AL TÍTULO DE
DOCTOR DE LA
UNIVERSIDAD DE BUENOS
AIRES, 2003**

**MECANISMO DE UNIÓN A ADN DEL DOMINIO C-TERMINAL DE LA
PROTEÍNA E2 DEL PAPILOMAVIRUS HUMANO, PRINCIPAL
REGULADOR DE LA EXPRESIÓN GÉNICA VIRAL.**

3638

La interacción proteína-ADN juega un papel central en la fisiología celular. Dentro de éstas, se identifica al reconocimiento secuencia específico con un rol regulatorio clave, pues de éstas interacciones depende la orquestación de los patrones de expresión génica de un organismo. Utilizando el dominio de unión a ADN del factor de transcripción E2 del papilomavirus humano, nos propusimos analizar la interacción en condiciones en las cuales fuera posible extraer datos fisicoquímicos precisos del proceso de unión. Confiamos en que la interpretación de éstos en términos termodinámicos provee un marco para relacionar las estructuras proteicas con su funcionalidad biológica.

Caracterizamos un mecanismo de unión E2c-ADN al equilibrio en el cual E2c presenta dos modos diferentes de unión de ADN diferentes. Estos modos resultan estar estrechamente relacionados con la ocurrencia de poblaciones conformacionales del estado plegado de E2c. Mostramos además que esto no sucede en un homólogo cercano con el que comparte una alta identidad de secuencia y una topología de plegamiento idéntica, sugiriendo que esto puede estar relacionado con los distintos roles funcionales que E2c presenta.

El camino cinético de interacción E2c-ADN se inicia con una colisión rápida, limitada por la difusión de las macromoléculas en solución. Esto da lugar a un "complejo de encuentro" estabilizado por interacciones electrostáticas, en el cual E2c tiene la posibilidad de deslizarse sobre la doble hélice. Si el ADN contiene un sitio E2, el complejo sufre fuertes rearrreglos conformacionales que, luego de una fase de expulsión de solvente, resultan en la formación de un complejo estable proteína:ADN. Este último paso involucra un ocultamiento de superficie accesible al solvente y es allí donde se consolidan las interacciones cercanas aminoácido-base. Experimentos de doble-mezcla por flujo detenido nos permitieron asignar el orden secuencial de los rearrreglos y demostrar que la unión puede también proceder por una ruta macroscópica paralela, caracterizada como de dos estados.

Utilizando una estrategia de mutagénesis sitio-dirigida realizamos un mapeo energético de la interfase de unión. Pudimos asignar las contribuciones que cada residuo realiza, en forma aditiva, a la energía libre de asociación. Las asignaciones energéticas resultaron estar en excelente concordancia con las estructuras de alta resolución, y correlacionan con el carácter altamente dinámico de la interfase de asociación.

Los conjuntos de estados de transición del camino cinético fueron analizados con 23 mutantes sitio-dirigidos. Esto nos permitió disectar las contribuciones energéticas de los residuos independientes a nivel atómico e inferir los cambios conformacionales que ocurren durante las etapas de unión y reconocimiento de secuencia. Mostramos que la plasticidad estructural que E2c presenta está íntimamente ligada a su capacidad de reconocimiento sitio-específico.

Por último, presentamos un modelo funcional plausible de regular la actividad biológica de E2c y su capacidad de reconocimiento de secuencias de ADN.

Protein-DNA interactions play a central role in cellular physiology. Within these, site-specific recognition is a primary topic that underprints the expression profile of whole genomes. Although many biological aspects of these non-covalent interactions are becoming clearer, detailed biophysical understanding of site specific recognition is lacking. Using the human papillomavirus (HPV) E2c transcription factor as a model system we aim a physicochemical dissection of its DNA binding properties. We set-up an spectroscopic binding assay, readily done in solution, which allowed us to quantify DNA binding parameters in energetic terms, detect conformational changes exerted on both macromolecules upon binding and relate them to the high resolution structures available.

We characterized an equilibrium DNA binding mechanism where the HPV-16 E2c domain is capable of two distinct binding modes, which correspond to different conformational ensembles of the E2c molecule. Despite having a strong sequence identity and an identical folding topology, a close homolog does not show this behaviour under the same experimental conditions. The differences observed may be related to the distinctive transcriptional roles in which these proteins are involved.

The kinetic DNA-binding mechanism is initiated by the diffusion controlled formation of an encounter complex stabilized by electrostatic interactions, in which E2c has the possibility to slide along the DNA duplex. If the target DNA contains an E2-binding site, substantial conformational rearrangements take place, and a solvent exclusion step leads to the formation of a highly stable protein-DNA complex. This last step involves the largest burial of surface area from the interface and the direct readout of the DNA bases is proposed to occur. Double-jump stopped flow experiments allowed us to characterize the sequence of intermediate events and demonstrate that a fast-formed final complex can be formed through a parallel two-state macroscopic route.

An extensive site-directed mutagenesis analysis over the entire DNA binding region led us assign the energetic contributions of individual sidechains to the overall binding free energy. We show that each sidechain contribute in a small amount to the total free energy in an additive way. These contributions are in excellent agreement with the high resolution structures and correlate with the highly dynamic character of the interface.

The transition state ensembles of the kinetic binding pathway were probed with 23 site-directed mutants. This led us to dissect the sidechain energetic contributions at almost atomic resolution, and infer the conformational changes that take place in each kinetic event. We show that the conformational plasticity this protein presents is intimately linked to its DNA recognition capacity.

Finally, a plausible biological model for functional E2c regulation and specific DNA recognition is presented.

AGRADECIMIENTOS

a quien seguiré llamando "jefe" por años, Gonzalo, por la profesional dedicación, la incondicional entrega y la interminable paciencia que tuvo que poner para poder enseñarle a un biólogo lo que significan "bioquímica de proteínas" y "rigor científico".

al professor doctor Goldbaum por compartirme sus consejos de estudio, su criterio y su humildad. A Don Wolosiuk por dejarme intuir lo que vive en la sutileza.

a mis compañeros del laboratorio de "Estructura-guión-Función": Leo, su precisión y sus brillantes aforismos; Ale, con la pregunta certera y la dulzura prestas; Juan, el empuje, la garra y los meses con las mutantes; Cristian, tipazo que destila sapiencia; María, que no se cansa de mostrarme que soy un inmaduro; Eleo, instantánea vecindad; Clarita, con su visión no-filatética de una proteína; Marianito, que se hace cargo estoicamente, Santlagos (aquel que cristallizaba proteínas de hígado de bagre, y éste que tomó la posta) y a Lili, con la voluntad siempre puesta y su contagiosa alegría

a la estimada Dra Cerutti, siempre con los ojos abiertos; a Pato, Julio y Andrea, y las discusiones proteicas por las cuales vagamos.

a los compañeros pipeteadores del instituto, los enriquecedores intercambios y los viernes de copetín. A radio pasillo, por informar fidedignamente de los acontecimientos locales.

al personal estable de la (ex) Fundación Campomar, especialmente a Eva, Laura, Dora y Darío, y a María Jimena Ortega, porque además de gritar juntos el GOLI con las secuencias, me muestra un mundo fuera del eppendorf.

a los amigos de la vida, cercanos a pesar de las distancias. A Gabie, por el diseño editorial y las horas, siempre insuficientes, con la máquina.

a "Soy Cuyano", que me presentó, entre otras cosas, el trabajo de H.M. Andrade.

a mis familias, que sé que estuvieron, están y estarán enseñándome y apoyándome durante toda mi vida.

a Lau, por todo lo demás.

El presente trabajo contó con el apoyo económico de: Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Universidad de Buenos Aires, Fundación Campomar (ahora Fundación Leloir), Fundación Antorchas y Wellcome Trust.

A mi viejo,
que me enseña a no traicionarme y a escuchar el corazón.

INTRODUCCION

EL MAESTRO:- ¿Hacia qué opinión se inclinó por fin la filosofía?

SI FU:- O bien las cosas están fuera de nosotros, por sí y aún sin nosotros, o bien las cosas están en nosotros, para nosotros y no sin nosotros.

EL MAESTRO:- ¿Cuál es la opinión correcta?

SI FU:- Todavía no se ha llegado a una decisión.

EL MAESTRO:- ¿Hacia qué opinión se inclina la mayoría de nuestros filósofos?

SI FU:- Las cosas están fuera de nosotros, por sí, y aún sin nosotros.

EL MAESTRO:- ¿Por qué quedó sin decidir la cuestión?

SI FU:- El congreso que debía decidirla se reunió hace doscientos años en el monasterio Mi Sang, que se encuentra a orillas del río Amarillo. La cuestión era ésta: ¿existe realmente el río Amarillo o existe sólo en la cabeza de la gente? Pero mientras se celebraba el congreso, se produjo un deshielo en las montañas, el río Amarillo se desbordó, inundó el convento de Mi Sang y se llevó a todos los participantes del congreso. De manera que no llegó a darse la prueba de que las cosas existan fuera de nosotros, por sí y aún sin nosotros.

Bertold Brecht

J. D. Watson comienza un difundido libro de texto sentenciando: "No hay sustancia tan importante como el ADN" ¹⁰⁷. Posición comprensible de parte de una de las personas que ha *inventado* y vivido de la famosa molécula helicoidal por más de 50 años. En esta sentencia no parece quedar implícito el rol funcional e Integrador que cumplen las proteínas sobre la expresión de la información genética, en gran parte codificada en el ADN. Son las proteínas las que proveen de *vitalidad* al químicamente inerte y "tan poco reactivo ADN" ⁵³.

La asociación entre estos grupos de macromoléculas, proteínas y ADN, se describió medio siglo años antes del establecimiento del ADN como "principio transformante" de *Pneumococos* en 1944³, con la co-purificación de las preparaciones proteicas de "nucleína" con una sustancia fosfoglicosídica, ADN. Lo que hoy conocemos por histonas fueron tal vez las primeras proteínas descritas que participan de las interacciones de las que trata la presente tesis.

Durante la emergencia intelectual de lo que hoy conocemos como "Biología Molecular" se establecieron los principios fundacionales que permitieron comenzar a estudiar los mecanismos por los cuales se regula la expresión génica celular. Las primeras ideas mecanísticas fueron presentadas por Francois Jacob y Jaques Monod, y resumidas en lo que se conoce como el modelo del *operón* ⁴³. Según este modelo existen "elementos reguladores" presentes *en cis* al ADN, que interactúan específicamente con elementos *en trans*, proteínas, regulando la expresión de uno u otro gen. Es la interacción entre el elemento *cis* (*lacO*, una secuencia de bases específica) y el elemento *trans* (LacR, la proteína represora) la que tiene un rol *decisivo* en la regulación de la expresión génica, ya que de esta interacción depende la expresión de las proteínas estructurales y regulatorias del metabolismo celular. Este modelo general de regulación, que involucra el reconocimiento específico proteína-ADN, se extiende hasta hoy en día cuando conocemos con exquisito detalle miles de elementos *cis* y *trans*, constituidos en redes de regulación que orquestan la expresión génica celular ^{11, 77}. Aún así, hay preguntas que subsisten desde los primeros años del desarrollo de los modelos mecanicistas: ¿Cómo se da el reconocimiento entre los elementos *cis* y *trans*? ¿Porqué el elemento *trans* X se une a elementos *cis* Z y no a otro/s?

INTERACCIÓN PROTEÍNA-ADN: LA ESPECIFICIDAD

Las interacciones proteína-ADN se identifican como un proceso central de la fisiología celular. Entre las funciones asignadas a éstas, se cuentan algunas en las cuales el reconocimiento es inespecífico, es decir, proteínas que unen ADN de cualquier secuencia de bases. Un ejemplo clásico lo constituye la interacción ADN-histonas, que si bien es no-específica, media procesos extremadamente relevantes en la regulación de la expresión génica, como lo evidencian los recientes avances en el campo de la epigenética²⁶. Las interacciones inespecíficas proteína-ADN participan también en procesos fisiológicos como la recombinación genética, la segregación cromosómica, la síntesis y "reparación" de ADN, etc.²⁹.

De las proteínas que interactúan con ADN, algunas poseen la extraordinaria capacidad de unirse selectivamente a una secuencia de bases en particular en el contexto de un genoma que presenta un gran exceso de potenciales blancos de interacción. Un modelo estudiado en profundidad es el de la interacción entre el LacR y *lacO*. Ya en la década de 1970, von Hippel mostró que la capacidad de selección de sitios puede explicarse por la afinidad diferencial entre la interacción LacR:*lacO* y LacR:ADN no-específico. Si se considera que existe un solo sitio *lacO* en todo el genoma de *E. coli*, hay también 5×10^6 sitios de la misma longitud que son "no específicos". Sin embargo se ha visto que, en una condición *in vitro*, ésta proteína presenta un "índice de discriminación" de 10^6 , definido esto como la afinidad de LacR por ADN específico relativa a la afinidad por cualquier ADN. Si estos índices se conservan *in vivo* se puede estimar que, si existe una sola molécula de LacR por bacteria, ésta pasará el 20% del tiempo unida a *lacO*. De forma análoga, sólo 20 moléculas de LacR bastarían para mantener el sitio *lacO* ocupado el 99% del tiempo.¹⁰⁵ ¿Cómo es posible que se de esta fantástica diferencia de afinidades?

ESPECIFICIDAD DE INTERACCIÓN: EL PROBLEMA BIOQUÍMICO-ESTRUCTURAL

La biología estructural es uno de los campos desde los cuales se han realizado fuertes aportes a la comprensión del reconocimiento proteína-ADN. En los últimos 20 años ha sobrevenido un virtual estallido en la elucidación de estructuras macromoleculares a alta resolución, entre las cuales contamos hoy con datos para 937 complejos proteína-ADN³². La convergencia de modelos estructurales, bioquímicos y genéticos, permite identificar las más diversas "estrategias químicas" naturales que median el reconocimiento específico entre proteínas y ADN.

A diferencia de la interacción proteína-proteína, en donde los pares interactuantes presentan diversos motivos estructurales con variadas características químicas, en la interacción proteína-ADN uno de los interactuantes es, a primera vista, químicamente uniforme. La elegante estructura de doble hélice¹⁰⁸, muestra al ADN de doble cadena como un polímero lineal uniforme, con una alta densidad de carga negativa distribuida a lo largo del esqueleto azú-

car-fosfato, encerrando un "core" parcialmente hidrofóbico formado por pares de bases apiladas a lo largo del eje central. A pesar de esta uniformidad general, se destacan asimetrías presentes en los surcos mayores y menores, donde los átomos laterales de las bases quedan expuestos hacia el exterior de la doble hélice, formando el *piso* de los surcos mayores y menores. Dado que cada par de bases tiene un conjunto característico de grupos funcionales expuestos, cada secuencia de ADN posee una potencial *firma química*, formada por el patrón de exposición de grupos funcionales en los surcos ³⁸. Hoy en día se acepta que, junto con las variaciones estructurales del clásico ADN en forma B (algunas de ellas secuencia-dependientes), es ésta superficie la que es reconocida específicamente por las proteínas. Incluso antes de la determinación estructural del primer complejo proteína-ADN, y basados en las dimensiones físicas de los elementos de estructura secundaria de las proteínas y el B-ADN, se postularon mecanismos de reconocimiento que involucran la inserción de motivos de hélices- α y hebras β -antiparalelas en los surcos, proveyendo un modelo estructural plausible para el reconocimiento de patrones específicos de contacto ¹⁴.

La resolución cristalográfica de complejos proteína-ADN confirmó algunas de las hipótesis estructurales de interacción. El consenso generalmente aceptado hoy en día es que las proteínas distinguen secuencias de bases por medio de la interacción de superficies complementarias al ADN, formando redes de contactos cercanos entre las bases y los residuos aminoacídicos que constituyen las superficies. A éste reconocimiento directo aminoácido-base se lo conoce como mecanismo de *lectura directa* de secuencia ⁴⁶. Sumado a los contactos íntimos mencionados, se ha identificado que todos los complejos proteína-ADN forman una gran cantidad de contactos salinos entre grupos aminoacídicos cargados positivamente y el esqueleto de azúcar-fosfatos del ADN. En principio, estos contactos son considerados como "no específicos", dado que el esqueleto exterior es común a todas las secuencias de ADN.

Influido tal vez por la simplicidad del código genético, muchos investigadores se empeñaron en definir reglas básicas de interacción aminoácido-base que pudieran explicar la especificidad de la interacción proteína-ADN en general ⁶⁷. Hoy es claro que tal *código simple* de reconocimiento no existe ⁶⁸, en tanto que el contacto aminoácido-base está fuertemente influido por la topología proteica general, las distorsiones estructurales de las macromoléculas, los múltiples eventos dinámicos asociados a la unión y los fuertes efectos del microentorno sobre las interacciones macromoleculares. Las variables necesarias para una descripción detallada basada en la complementariedad de las bases y aminoácidos son prácticamente infinitas, y hoy resulta claro que el problema no es reducible a mínimas interacciones. Sumado a esto, se reconoce que existen efectos sobre la interacción específica que están mediados por bases y aminoácidos que no se contactan directamente, mecanismo que recibe el nombre de *lectura indirecta* de secuencia.

ESTRUCTURAS DE ADN UNIDO A PROTEÍNAS

La longitud efectiva de la secuencia *operadora* de ADN está estrechamente relacionada con el tamaño y la conformación de la proteína que la reconoce. En general, los sitios de reconocimiento específico están contenidos entre 6 y 18 pb. En muchos casos las secuencias de los sitios son palíndromos, reflejando la existencia de simetrías de contacto presentes en la proteína respectiva, en general dimérica o pseudodimérica. La repetición sucesiva de los mismos contactos duplica la energía libre de interacción, y por tanto aumenta la afinidad específica en órdenes de magnitud, ya que se relacionan exponencialmente según: $\Delta G = -RT \ln K_D$.

En gran parte de las estructuras cristalográficas resueltas para complejos proteína-ADN se identifica al ADN en forma-B, con pequeñas distorsiones en parámetros relacionados con los ángulos que forman los planos de las bases y, algunos casos, algún grado de curvatura del eje mayor del ADN. En una baja proporción, la estructura del ADN se encuentra fuertemente deformada. El ejemplo más drástico es el de la interacción de TBP ("TATA binding protein"), en la cual el ADN presenta una curvatura de más de 90° sobre el eje central. A pesar del alto costo energético de esta deformación, éste es compensado por un gran número de contactos favorables con la proteína, asociados además a los efectos de remoción de solvente de la interfase de unión, y otros efectos sobre el entorno proteico^{61, 62}. Aunque a primera vista estas distorsiones estructurales pueden parecer irrelevantes, se ha identificado que pueden constituir hechos funcionalmente significativos por los efectos que puede ejercer sobre la regulación génica mediada por el contacto distante con otras proteínas⁶³.

ESTRUCTURAS DE LAS PROTEÍNAS QUE INTERACTÚAN CON ADN

Las proteínas que unen ADN muestran una gran variedad topológica. Las primeras tres estructuras resueltas, el represor λ ⁷⁹, λ -cro⁷⁶, y CAP ("Catabolic Activator Protein")⁷⁰, contienen un motivo estructural común, llamado hélice-vuelta-hélice, característico de proteínas de unión a ADN. Poco tiempo después se identificaron más arreglos conformacionales que revelaron las más diversas adaptaciones estructurales mediando el reconocimiento secuencia-específico. Las proteínas que unen ADN se clasifican hoy en más de 30 familias, con diferentes topologías de plegamiento y una gran variedad de elementos de estructura secundaria mediando el contacto directo⁶¹. Algunas topologías resultan comunes a varios tipos celulares, reconociéndose en ellas motivos prototípicos como "dedos de zinc", *homeodomain*, bZIP, etc., y otros menos frecuentes y restringidos a un grupo de organismos en particular, como el barril β -dimérico⁶².

En general, los plegamientos típicos poseen un dominio globular cuyas cadenas laterales de superficie interactúan directamente con ADN. Sin embargo, varios motivos de unión contienen zonas flexibles o desordenadas que se pueden identificar como estructuradas luego de la unión con ADN. En los últimos años se ha identificado que existe un acoplamiento estre-

cho entre el proceso de *plegamiento* proteico y la unión de ADN ^{86,91}. Incluso se ha propuesto que proteínas que en un principio puede parecer estructuradas, sufren procesos de desplegamiento y plegamiento local que contribuyen significativamente al proceso de reconocimiento secuencia-específico ⁸⁵. Este punto se discutirá más adelante en relación con los resultados aquí presentados.

MODELOS GENERALES DE RECONOCIMIENTO PROTEÍNA-ADN

Para que una unión sitio-específica sea biológicamente relevante, la proteína tiene que unirse significativamente más fuerte a los elementos de respuesta (i.e. sitios específicos) que a otros sitios de ADN presentes en el genoma. Las uniones a sitios inespecíficos generalmente se encuentran en el rango de 10^3 - 10^6 M ($\Delta G \sim -4$ a -7 kcal/mol). En el lado opuesto de la escala, la afinidad de la unión por sitios específicos no tiene que ser excesivamente alta para que la unión sea fisiológicamente reversible y el sistema no resulte "hipersensible" a pequeñas fluctuaciones la concentración de proteína endógena ¹⁰⁵. Teniendo en cuenta las concentraciones intracelulares de proteínas regulatorias, sitios específicos y ADN total, se estimó que estos procesos tendrían sentido fisiológico para constantes de asociación con un límite superior de 10^{12} M ($\Delta G \sim -16$ kcal/mol) ⁵⁷.

La disección de las contribuciones termodinámicas a la energía libre de interacción puede proveer información sobre los mecanismos moleculares que operan en la interacción ⁴⁷. La termodinámica clásica muestra que la energía libre (ΔG) tiene componentes entálpicos (ΔH) y entrópicos (ΔS) relacionados por: $\Delta G = \Delta H - T\Delta S$. Es de esperar que la formación de los contactos cercanos de reconocimiento (uniones de hidrógeno, pares iónicos, y contactos no-polares) sean entálpicamente favorables ($\Delta H < 0$). Oponiéndose a este efecto existe una entalpía desfavorable de desolvatación de grupos, ya que a medida que forma la interfase consolidada, las superficies polares y no-polares se acercan. La desolvatación de las superficies polares es una fuente de ΔH desfavorable. La desolvatación de las superficies no-polares contribuye en poca medida al ΔH a 298°K (sin embargo es la fuente mayoritaria de ΔS favorable). Otro de los aportes al ΔH desfavorable es el llamado componente de *tensión* ("strain"). Este ocurre cuando átomos, grupos funcionales o hasta residuos enteros (bases y aminoácidos) adoptan posiciones en el complejo proteína-ADN que no coinciden con el propio mínimo de energía potencial de la interacción local. Esto puede darse como resultado de torsiones de enlaces, rotaciones locales o energías de no-unión ("clash" estérico, repulsión electrostática, etc). Por supuesto que el resultado neto de las interacciones es minimizar la energía libre del complejo como un todo y no la energía de las interacciones locales. En cuanto a los componentes entrópicos, el ΔS favorable es contribuido mayoritariamente por la liberación de agua de las superficies hidrofóbicas y en menor medida por la desolvatación de superficies polares y por la redistribución de

contraiones. En oposición, las fuentes mayoritarias a las que se puede atribuir un ΔS desfavorable son la pérdida de grados de libertad de translación y rotación de las macromoléculas, las diversas restricciones conformacionales de los elementos unidos y las restricciones vibraciones de las moléculas de agua atrapadas en la interfase ⁴⁷.

A partir de los estudios realizados sobre los pocos modelos estudiados en profundidad (LacR, TBP, c-Myb), se podría generalizar un mecanismo simple de reconocimiento proteína-ADN de dos etapas ⁷⁵: 1) En la primera etapa, atracciones electrostáticas de largo alcance acercan los componentes. Los contraiones condensados sobre las macromoléculas son liberados al solvente a medida que se forman enlaces salinos entre los residuos cargados positivamente y los grupos fosfato. La unión inespecífica terminaría en esta etapa y estaría predominantemente guiada por un aumento entrópico favorable, mayoritariamente debido a la liberación de contraiones. Esta etapa está caracterizada por una afinidad relativamente baja y constantes de disociación cinética altas (disociación rápida). En esta etapa, las proteínas sitio-específicas interactuarían con el ADN independientemente de su secuencia de bases. 2) Cuando la proteína "encuentra" un sitio específico puede tener lugar la formación de contactos cercanos entre los aminoácidos y las bases. Esto daría lugar a la formación de un complejo de alta afinidad, caracterizado por constantes de disociación cinéticas bajas. Este evento estaría guiado entálpicamente por la formación de los contactos cercanos y entrópicamente por la liberación de solvente de la interfase de unión. Además, durante ambas etapas pueden ocurrir cambios conformacionales locales o globales, tanto en la proteína como en el ADN, que pueden contribuir a reforzar o debilitar la interacción ⁴⁷.

Resulta extremadamente difícil cuantificar desde primeros principios cada una de las contribuciones energéticas descritas para una Interacción proteína-ADN. Estas Interacciones involucran miles de átomos participando simultáneamente en Interacciones no-covalentes de diferentes características y variadas magnitudes, que además se presentan con modos vibracionales locales que, presumiblemente, no son independientes de los modos vibracionales globales. Con el fin de aportar a la comprensión general de los mecanismos de unión/reconocimiento proteína-ADN, por medio de la disección de las contribuciones energéticas de interacción, iniciamos el presente estudio energético-funcional utilizando como modelo la interacción entre el dominio C-terminal de la proteína E2 del *papilomavirus* humano (cepa16) y el sitio específico de ADN que esta proteína reconoce.

INTRODUCCIÓN AL MODELO UTILIZADO: E2C DE *PAPILOMAVIRUS*

PAPILOMAVIRUS: GENERALIDADES Y ASPECTOS CLÍNICOS

Los *papilomavirus* fueron identificados hace casi cien años, cuando en 1907 Cluff demostró la posibilidad de formación de verrugas cutáneas inducidas por el filtrado, libre de células, de un exudado de verrugas comunes¹⁵. El primer *papilomavirus* fue descrito por Richard Shope en 1933, que reconoció al CRPV (Cottontail Rabbit Papilloma-virus) como el agente etiológico responsable de la transmisión de papilomatosis cutánea de conejos¹⁶. Durante el siglo XX estos virus fueron refractarios al estudio con técnicas virológicas estándar, al no encontrarse una forma de propagar los *papilomavirus* en cultivos celulares, dificultad que persiste hasta el día de la fecha. La caracterización biológica y bioquímica de estos virus retoma fuerza recién en la década de 1970 con el advenimiento de las técnicas de biología molecular, que otorgaron la posibilidad de clonar los genomas virales y, durante los 80, con el desarrollo de ensayos de transformación de cultivos celulares *in vitro*. Hoy en día se distinguen *papilomavirus* infectivos de una amplia gama de mamíferos y, particularmente, más de 100 cepas de *papilomavirus* que infectan *H. sapiens*, los llamados HPV (por Human Papilloma Virus)¹⁷. Estudios evolutivos recientes muestran que los *papilomavirus* comparten cerca de 100 millones de años de coevolución con los mamíferos, por lo cual representan también un excelente sistema de estudio de relaciones coevolutivas a nivel genético-molecular.

Los *papilomavirus* fueron incluidos junto con los *Polyovirus* y los *Simian Vacuolating Virus* dentro de la familia de los Papovavirus, con los cuales comparten características biológicas, bioquímicas y moleculares comunes. Todos estos son virus oncogénicos, relativamente pequeños, con un genoma circular de ADN de doble cadena. Los *papilomavirus* infectan queratinocitos de epidermis mucosas o epidermis cutáneas estratificadas. Estas infecciones pueden derivar en dolencias varias, desde la formación de lesiones proliferativas benignas (verrugas comunes) hasta diversos carcinomas, siendo el más frecuente el carcinoma de cuello de útero. A pesar de su gran prevalencia de infección y la potencialidad de generar patologías severas, no existen aún terapias específicas desarrolladas para estos virus¹⁸.

Dentro de los *papilomavirus* humanos, las distintas cepas se encuentran asociadas a diferentes lesiones. Las proliferaciones cutáneas son causadas por las cepas de HPV-1, 2, 3, 4 y 10. Estas son lesiones benignas que presentan un crecimiento autolimitado, no invasivo, y forman *papilomas* o verrugas comunes en las extremidades corporales. La incidencia de éstas en la población mundial se estima entre el 7 y el 10%, siendo en general inócuas y revierten espontáneamente. Un segundo grupo lo componen las cepas HPV-5 y 8, asociados a *epidermodisplasia verruciformis* (EV). Los pacientes EV son sensibles al desarrollo de verrugas cutáneas que, ante la exposición a la luz solar, poseen alto riesgo de degenerar en transformación maligna¹⁹.

Los *papilomavirus* asociados a lesiones genitales (*condiloma acuminata*) son de una relevancia médico-sanitaria sustancialmente mayor. En estos casos, el HPV se transmite sexualmente generando a razón de un millón de nuevos casos por año, sólo en EEUU. El HPV es por tanto el virus de transmisión sexual más común, estimándose que entre el 5 y el 20% de la

población mundial entre 15 y 49 años es portadora de HPV ³⁹. Las lesiones anogenitales que estos virus desencadenan pueden desarrollar malignidades severas, siendo la más común de ellas el carcinoma de cuello de útero. El cáncer cervical es el segundo carcinoma más frecuente en la población femenina y el séptimo más frecuente de la población mundial, siendo responsable de una alta tasa de mortalidad, estimada en 200.000 personas anualmente¹⁹.

Los *papilomavirus* genitales son a su vez clasificados como de *alto* o *bajo* riesgo, según su asociación con la posibilidad de progresión maligna. Los HPV de las cepas 6 y 11 se encuentran en más del 90% de las lesiones genitales benignas, que pueden revertir espontáneamente o persistir por años sin causar mayor efecto, por lo que son considerados de *bajo* riesgo. Estas dos cepas son también asociadas con *papilomatosis laríngea*, un tumor benigno del epitelio de la laringe. Las cepas de HPV 16, 18, 31, 33 y 35 son las consideradas como de *alto* riesgo. Estas pueden encontrarse en displasias moderadas a severas y en carcinomas invasivos de la mucosa genital. Si bien los papilomavirus de las cepas de *alto* riesgo no pueden considerarse como la única causa de la generación de tumores, estudios recientes muestran que prácticamente todos (99,7%) los carcinomas de cervix son positivos a la presencia HPV. De éstos, el 50% pertenecen a la cepa HPV-16, objeto del presente estudio ¹⁹. Como dato curioso, cabe mencionar que muchas de las líneas celulares comúnmente utilizadas en laboratorios de investigación biomédica son transformantes de HPV, caso HeLa (HPV18), SiHa (HPV16) y CaSki (HPV16) ¹⁰³.

A pesar que el HPV esta asociado a un amplio espectro de la población y a dolencias extendidas desde verrugas benignas a cánceres metastásicos, en la actualidad no existen tratamientos virus-específicos, como tampoco la posibilidad de vacunación efectiva. El tratamiento actual se desarrolla en forma no-específica por medio de la destrucción local o remoción del tejido infectado. En algunos casos es posible el tratamiento con la aplicación tópica de agentes citotóxicos: ácido siálico, TFA, podofilina, colchicina, bleomicina, etc. Desde ya que, dada la morbilidad y el alto costo económico y social producido por los HPV, el desarrollo de vacunas terapéuticas y profilácticas se encuentra en plena expansión ⁸².

CICLO DE VIDA Y ORGANIZACION GÉNICA

Los *papilomavirus* son altamente específicos de especie y de tejido. Como se mencionó, infectan células epiteliales y la expresión génica viral esta íntimamente ligada al estado de diferenciación de la célula infectada. En general, se acepta que la infección celular se da luego de un microtrauma en el tejido dérmico. Aún no se identificaron receptores celulares específicos del virus, aunque se ha reportado que la adsorción viral esta ligada a la Integrina alfa-6 ⁵⁹. Luego de la entrada y el desensamblado de la partícula viral, el genoma viral se movi- liza al núcleo celular, pudiendo permanecer como episoma libre o integrado al genoma de la

célula huésped. En el primer caso, la cantidad de copias se mantiene constante entre 20 y 100 genomas por célula, que replican en sincronía con el ADN celular. La expresión génica de los HPV altera drásticamente la diferenciación celular, siendo el efecto más común la estimulación de la proliferación celular ⁵⁹.

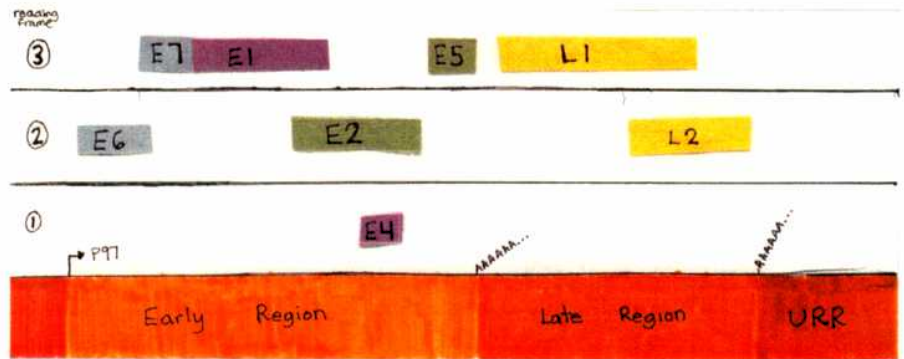


Figura 1.1: Esquema genómico de HPV-16. Se muestra un mapa extendido del genoma circular de HPV16. Los promotores y el origen de replicación se localizan en el URR, donde están los elementos de respuesta E2. En la parte superior se esquematiza la disposición de los 8 *orf* en los diferentes marcos de lectura. (adaptada de <http://www.stanford.edu/group/virus/papova>)

El genoma de los *papilomavirus* es de ADN doble cadena, circular, de un tamaño aproximado de 8 Kpb (Fig 1.1). En él se distinguen 8 marcos de lectura (*orf*), ubicados todos en la misma hebra. Estos se nombraron según el momento de expresión respecto a la replicación viral como tempranos ("early", E) o tardíos ("late", L). La secuencia completa del genoma del HPV-16 se publicó en 1985 ⁹⁰, donde se distingue, además de los *orf*, una región de aproximadamente 800pb denominada LCR (por "Long Control Region"), o URR (por "Upstream Regulatory Region"). En esta región se identificaron los elementos controladores de la replicación viral (*ori*), y la mayoría de los elementos regulatorios de la expresión génica ⁵⁹.

BREVE DESCRIPCIÓN DE LOS ORF DE HPV

L1 y L2: estos codifican las proteínas de la cápside viral, y su expresión se detecta exclusivamente en queratinocitos diferenciados ⁵⁹. L1 es la proteína mayoritaria de la cápside, un polipéptido de 55 KDa cuya secuencia esta altamente conservada entre las diferentes cepas. L1 es capaz de formar VLPs ("virus-like particles"), que inhiben la adsorción viral a células epiteliales en cultivo. L2 es una proteína de 75 KDa, componente minoritario de la cápside viral. En conjunción con L1 puede formar VLP que presentan una estequiometría L1:L2 de 30:1. La disposición de las 72 unidades morfológicas de la partícula viral (capsómeros) forman un icosaedro de simetría T=7. Los capsómeros de 12 vértices son pentavalentes y los restantes 60, hexavalentes, conformando una partícula viral madura de 55nm de diámetro. (Fig 1.2)

E5: el *orf* de E5 codifica un polipéptido de 40 aa, sin actividad enzimática descrita. La expresión de E5 se localiza asociada a membranas intracelulares y al aparato de Golgi. Se ha reportado que la expresión de E5 es capaz de transformar células de ratón en cultivo, estimulando el crecimiento celular. Como blancos celulares de interacción se identificaron al receptor de PDGF- β , el receptor de EFG, y la ATPasa vacuolar. Si bien su efecto oncogénico está en discusión, se acepta que E5 coopera con E6 y E7 en la estimulación de la proliferación celular ⁵⁹.

E6: junto con E7, este *orf* fue caracterizado como uno de los principales agentes responsables de la transformación celular. Como muchas proteínas virales que intervienen en el ciclo celular, la proteína E6 del *papillomavirus* puede formar complejos con p53. En células infectadas con HPV, los niveles de p53 son bajos, y se ha propuesto un mecanismo de degradación de p53 dependiente de ubiquitinación mediada por E6. Se ha reportado también que E6 interactúa con otros blancos celulares como paxilina, hDLG, E6AP, IRF-3, Bak, CBP/p300, etc. ⁵⁹

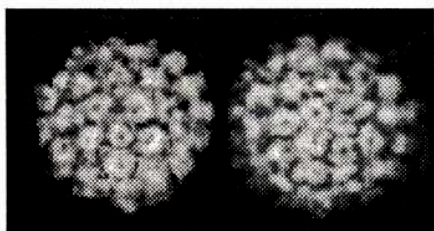


Figura 1.2: Micrografía de barrido electrónico de partículas virales de HPV. La partícula viral esta compuesta por 72 unidades morfológicas, formando una estructura icosaédrica de 55nm de diámetro ⁵⁹.

E7: el *orf* E7 codifica una proteína pequeña de alrededor de 100 aminoácidos, de localización principalmente nuclear. E7 es capaz de immortalizar líneas celulares en cultivo, tanto sola o en cooperación con el oncogen Ras o HPV-E6. Se identificó que E7 utiliza como proteínas blanco al supresor de tumores retinoblastoma (Rb) así como otras "pocket proteins" que incluyen p107 y p130. Es además capaz de formar complejos con el factor de transcripción celular E2F, y ciclinaA, todas estas involucradas en la regulación del ciclo celular. La función de E7 esta regulada por múltiples modificaciones post-traduccionales, incluyendo fosforilación de varios residuos y regulación redox ⁵⁹. Se ha reportado que numerosas propiedades bioquímicas de E7 son relevantes para su funcionalidad *in vivo*. Por ejemplo, E7 tiene características que la identifican como una proteína "nativamente desplegada", donde su funcionalidad esta acoplada al estado de plegamiento del polipéptido, a su vez asociado a la interacción con múltiples blancos celulares ¹.

E4: El producto del *orf* E4 se expresa como un componente tardío, si bien antes de la expresión de L1 y L2, con un rol poco claro en la infección viral productiva. Se ha reportado que la expresión de E4 no es esencial para la replicación ni la transformación celular ⁹⁹. Se la describió como asociada al citoesqueleto y se propone que está asociada al colapso de la red de citoqueratina, pudiendo contribuir durante la replicación vegetativa del virus. Como dato curioso, nótese que el *orf* E4 se superpone con el de E2, en otro marco de lectura (Fig 1.1).

E1: es la única proteína codificada en el genoma viral que presenta actividad enzimática, siendo un factor esencial para la replicación del genoma viral. Por la cercana homología que presenta con el antígeno-T de SV40, se infirió que E1 debería estar involucrado en una función análoga en los *papilomavirus*, como iniciador de la replicación viral. Este *orf* se expresa en muy bajas cantidades y la proteína derivada presenta funciones enzimáticas de ATPasa, unión de nucleótidos y helicasa ATP dependiente. Es una fosfoproteína de 68KDa, de localización nuclear, que presenta unión específica, pero débil, al origen de replicación. E1 es requerida para iniciación y la elongación de la síntesis de ADN y se la ha descrito interactuando con varias proteínas celulares: subunidad p180 de la DNAPol α -primasa, Histona H1, SW1/SNF5, ciclina E/Cdk2, Hsp40/Hsp70 y Ubc9 ¹¹⁰. Se identificó además que la interacción específica de E1 con la proteína viral E2 es fundamental para iniciar la replicación viral ⁹⁹ (ver más abajo).

E2: al producto del *orf* E2 se lo distingue involucrado en la regulación de la transcripción de todos los genes virales y en el inicio de la replicación episomal. Fue identificado por primera vez como un activador de la transcripción ⁹⁴, capaz de regular la transcripción por medio de elementos de respuesta (E2-RE) localizados en *cis* al genoma viral ⁹³. El *orf* E2 presenta alta similitud entre los *papilomavirus* y se pueden identificar en él tres dominios conservados: unos 150 aminoácidos N-terminales, conectados por una región flexible de aprox 100 aa, con otro dominio C-terminal de alrededor de 100 aminoácidos (Fig 1.3). Al dominio N-terminal se lo identificó involucrado en la transactivación a partir de secuencias promotoras. El dominio C-terminal, objeto del presente trabajo, es el responsable de mediar tanto la homodimerización de E2 como la capacidad de unión de ADN en forma secuencia específica. La proteína E2 del *papilomavirus* bovino (BPV-1) fue la utilizada como primer modelo de estudio de la proteína E2. Para ésta se identificaron 3 especies proteicas, producto de *splicing* alternativo: la traducción del *orf* completo (E2-TA, por transcriptional activator), que puede funcionar como activador o represor de la transcripción, dependiendo de la localización del sitio de unión respecto al promotor o a los *enhancer* ⁶⁸. Las otras dos formas, más pequeñas, llamadas E2-TR y E2-E8, fueron descritas como represores de la transcripción. Estas dos últimas carecen del dominio N-terminal y se las supone compitiendo con E2-TA por los sitios de unión a ADN ³³ (Fig. 1.3).

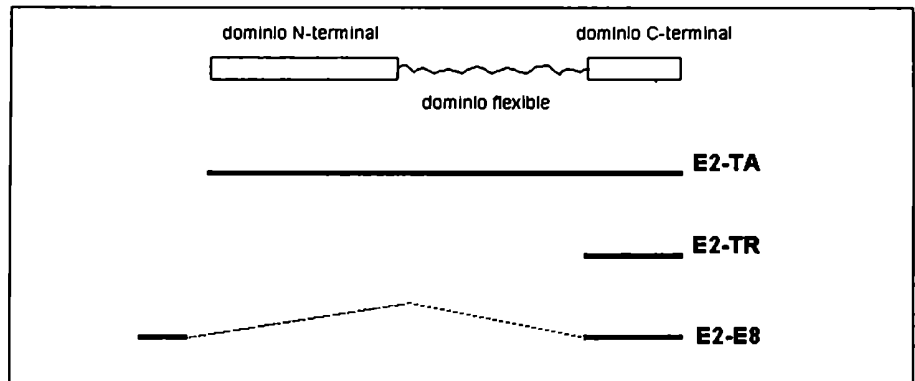


Figura 1.3: Esquema del *orf* E2. Se muestran los 3 dominios identificados en E2 en forma esquemática, N-terminal de trasactivación, el dominio flexible o "bisagra" y el C-terminal, de dimerización y unión a ADN (E2c). Abajo, los mRNA derivados de *splicing* alternativo.

La inactivación del *orf* E2 tiene efectos pleiotrópicos sobre la fisiología viral, disruptiendo la transformación celular, la replicación viral, y las funciones asociadas a la regulación de la transcripción. Se ha demostrado que la expresión de todos los genes tempranos del *papilomavirus* esta bajo el control de E2, mediada por la interacción de ésta con sitios específicos localizados en la región de control, URR (Fig 1.1). La secuencia consenso de los sitios E2 es un palíndromo de 12 pb: ACCGnnnnCGGT. En el caso de BPV, se identificaron 17 sitios E2, de los cuales 12 están en el URR. En el caso de los HPV, se identifican 4 sitios en el URR, cuya localización relativa está conservada entre las diferentes cepas. Los sitios 35 y 50 se encuentran río arriba del promotor P97, el cual que rige la expresión de los *orf* tempranos ⁴². La actividad de transcripción basal depende de los *enhancers* (queratinocito específicos) localizados en el URR. Por otro lado, la represión de la expresión génica mediada por E2 está también relacionada con la unión específica a los sitios 35 y 50, donde podría interferir en el ensamblado de la maquinaria de transcripción basal ⁴².

En el contexto fisiológico de una infección de *papilomavirus*, E2 es una proteína multifuncional. Puede participar como activador de la transcripción, como represor, como iniciador de la replicación, con funciones en la segregación episómica, e incluso se lo reportó como asociado al *splicing* alternativo de mRNAs de HPV ³¹. Todas estas funciones están mediadas por la interacción de E2 con otras proteínas, tanto virales como celulares. Dentro de los blancos celulares de interacción de E2 se identificaron varias proteínas que están involucradas en los procesos mencionados: TATA binding protein (TBP), TFIIB, AMF-1/Gps2, YY1, y más recientemente identificados, TopBP1, RACK1, P27BBP, POMP, ODC, Delta-Adaptina y PME-1 ^{9,19}.

E2 Y SU ROL EN LA PROLIFERACIÓN CELULAR

Hay por los menos dos vías por las cuales E2 puede afectar la proliferación celular. Por un lado influye indirectamente al afectar la expresión de los oncogenes E6 y E7. Se ha visto que células HPV positivas presentan cambios en los niveles de expresión de E2 que pueden resultar tanto en un aumento como en una disminución de la expresión de E6 y E7. La expresión de E2 (de BVP1) induce un arresto del ciclo celular en líneas derivadas de carcinoma de útero (positivas para HPV). En estas células, E2 reprime la transcripción de E6 y E7 y esto a su vez bloquea la transición de G1 a S ¹⁹. Por otro lado, se ha descrito que la expresión de E2 (de HPV-16, 18, 31), induce fenómenos de apoptosis, incluso en células que son HPV negativas. Asimismo, se reportó que la expresión de E2 (HPV16) puede influir la proliferación de células infectadas con HPV, independientemente de la expresión de E6 y E7 ¹⁹. El hecho de que mutantes de E2, defectivas en la capacidad de unir ADN, pueden afectar la proliferación y muerte celular, apunta a que este efecto puede estar mediado por la interacción de E2 con otros blancos celulares, y no por su capacidad de unión a ADN. Recientemente se ha descrito que E2 puede interactuar con p53 y con CBP ⁶⁵, un co-activador del proceso apoptótico independiente de p53. Por lo descrito, E2 de HPVs de *alto riesgo* parece actuar como un supresor de tumores. La disrupción del *ori* de E2 o la pérdida de su expresión, deviene en células que son más propensas a proliferar y, como consecuencia, más propensas a la generación de lesiones tumorales.

E2 COMO OBP

La replicación del genoma viral *in vivo* requiere de la proteína E2, aunque ésta no es necesaria para iniciar la replicación *in vitro*. Las proteínas iniciadoras de la replicación (*Origin Binding Proteins*, OBP) se unen a los orígenes de replicación (*ori*) y orquestan el ensamblado de la maquinaria de replicación de ADN. Si bien las actividades de las OBPs varían, todas comparten el rol de unirse y "marcar" el origen de replicación. En este sentido, deben ser capaces de unir específicamente el *ori*, en el contexto de un gran exceso de ADN inespecífico. En el caso de los *papilomavirus*, E1 fue identificado como el factor iniciador pero, contrariamente a lo esperado, E1 presenta baja capacidad de selección de sitios específicos. Es aquí donde E2 es requerida, dado que la selectividad de sitios de E1 se incrementa cientos de veces por la interacción con E2. Se ha descrito que estas proteínas unen sitios de ADN adyacentes y cooperan en la unión específica, formando complejos macromoleculares capaces de una alta discriminación sitio-específica. En un modelo reciente se propone que la formación de estos complejos se inicia con el ensamblado de un complejo de dos dímeros de E2 y dos dímeros de E1, de los cuales E2 es luego desplazada (de forma ATP dependiente) a medida que se forma la unidad funcional hexamérica de E1 ⁶⁶.

En los últimos años se le han asignado otras actividades a E2, por ejemplo las relacionadas con el mantenimiento y segregación del genoma viral ⁴². Se ha visto que plásmidos que contienen el origen de replicación de HPV son capaces de replicación transiente en células que expresan E1 y E2, aunque éstos se pierden al cabo de pocas generaciones celulares. El mantenimiento del plásmido a largo término requiere de elementos cromosómicos en *cis* que consisten de múltiples copias de sitios E2 ⁴³. Junto a la observación de que E2 se asocia a cromosomas mitóticos ⁴, esto sugiere que E2 podría estar involucrada en un mecanismo de acoplamiento entre la asociación del genoma viral a la cromatina y la replicación del genoma viral. Esto sería particularmente relevante durante la fase de latencia de la infección de HPV.

EL DOMINIO C-TERMINAL DE E2 (E2c)

Todas las funciones de E2 descritas en las secciones anteriores están mediadas por la capacidad de unir específicamente secuencias de ADN. A fines de la década de 1980 se demostró que los ~80 aminoácidos C-terminales de la proteína E2 conforman un dominio de plegamiento independiente ⁴⁴. Se describió además que éste es capaz de formar homodímeros que retienen la capacidad de unir específicamente ADN ^{45, 46}. A lo largo del presente trabajo se referirá a este dominio como E2c. Estudios iniciales sobre E2c, basados en comparaciones de secuencia primaria y en las topologías proteicas descritas como de unión a ADN, hipotetizaron que E2c podría tener un plegamiento clásico: cierre de leucinas, o hélice-vuelta-hélice ⁴⁴. Resultó sorprendente cuando en 1992 el grupo de Paul Sigler resolvió la primera estructura cristalográfica de un dominio E2c unido a ADN, mostrando una topología de plegamiento completamente novedosa: el barril β -dimérico ³⁴ (Fig 1.4). Hoy en día se considera a E2c como prototipo de una nueva clase estructural de proteínas que unen ADN, topología compartida con el antígeno nuclear del virus Epstein-Barr (EBNA1), virus con el cual HPV carece de relación filogenética cercana ⁷. La topología secundaria del barril β -dimérico es " $\beta 1-\alpha 1-\beta 2$ -loop- $\beta 3-\alpha 2-\beta 4$ ". Cuando la dimerización, los residuos de las hebras $\beta 2$ y $\beta 4$ de los "bordes" de cada subunidad forman una red continua de uniones de hidrógeno, que resulta en la formación de un barril de 8 cintas β -antiparalelas (Fig 1.4). La interfase del dímero es extensa y sustancialmente hidrofóbica, ocultando entre 1500 y 2500 Å² según las estructuras de E2c resueltas hasta la fecha ³⁵. Por fuera del barril- β , cada monómero expone dos α -hélices empaquetadas sobre la parte externa del barril- β , y se identifica a la mayor de ellas (hélice1) como involucrada en el contacto base específico (Fig 1.4). La hélice1 es por esto también llamada "hélice de reconocimiento". En todas las estructuras cristalográficas resueltas para complejos E2c-ADN, se distinguen a las hélices de reconocimiento insertadas en sucesivos surcos mayores de la doble hélice de B-ADN, contactando directamente motivos específicos de las bases nitrogenadas ^{34, 35}. Se destaca además una deformación del B-ADN, particularmente una curvatura de entre 43° y 51°, centrada en la zona del "separador" de los hemisitos.

La secuencia primaria de los dominios E2c de las diferentes cepas de *papillomavirus* está altamente conservada. Teniendo en cuenta la secuencia toda (~80aa), la conservación varía desde un 54% de identidad (77% similitud) entre los *papillomavirus* cercanos (HPV16, HPV18), hasta 33% de identidad (51% similitud) para los más distantes (HPV16 - BPV1). Los valores mencionados aumentan considerablemente si se evalúa sólo la zona correspondiente a la hélice de reconocimiento.

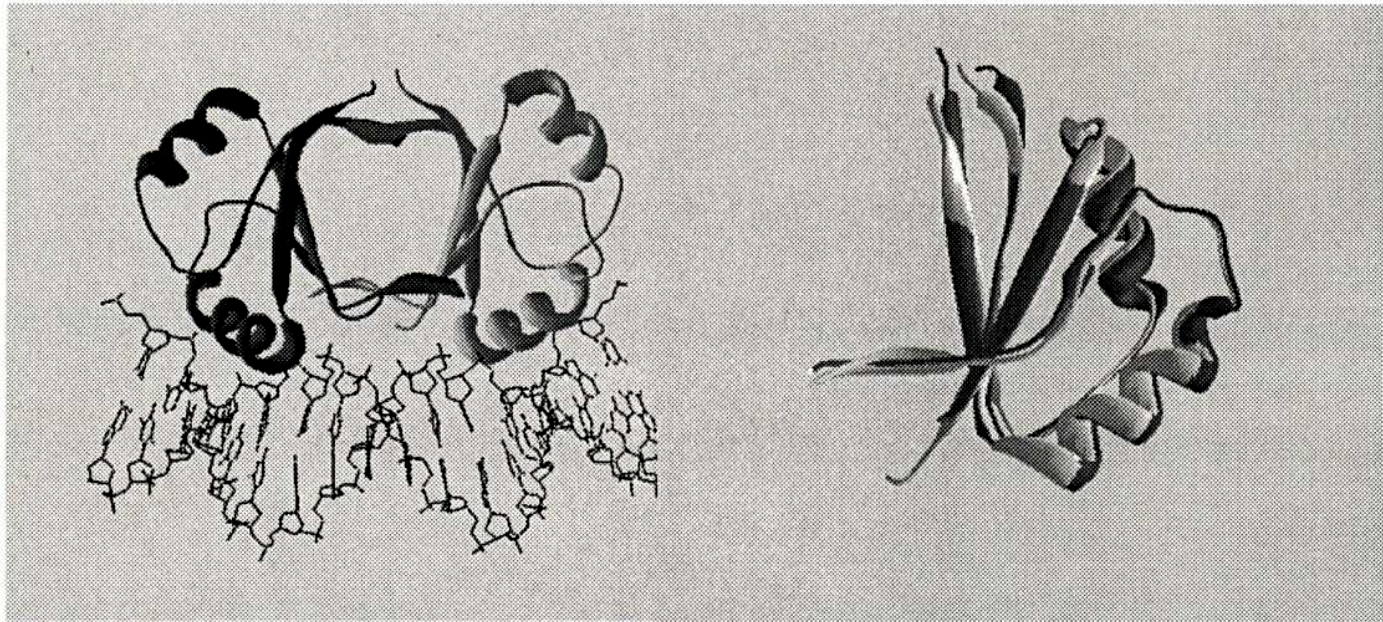


Figura 1.4: Estructura del dominio C-terminal de E2. E2c presenta una topología de barril β -dimérico, donde cada monómero aporta 4 cintas- β antiparalelas y dos α -hélices.
 izq. Modelo derivado del co-cristal de HPV18-E2c:ADN. Se muestran los monómeros con diferente tonalidad.
 der. Superposición de un monómero de E2c de HPV16 y HPV18.

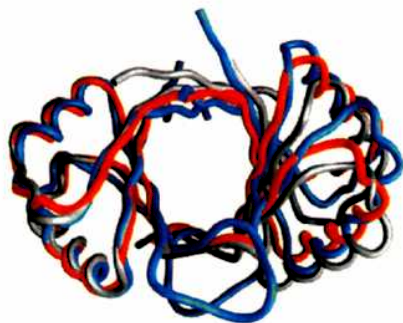


Figura 1.5: Dominios E2c de diferentes cepas de papilomavirus. Se superpusieron las estructuras resueltas de E2c de HPV16 (rojo), HPV31 (azul) y BPV1 (gris). Las diferencias cuaternarias se evidencian al superponer el monómero de la izquierda y dejar libre la orientación del otro monómero.

Hoy en día contamos con estructuras de alta resolución para los dominios E2c de varios *papilomavirus*. Cristalográficamente se han resuelto las de BPV1, HPV16^{36, 37}, y en co-cristales con ADN las de BPV1 y HPV18^{34, 48}. Por RMN se han resuelto las de BPV1⁵⁴, HPV31¹⁰⁴ y más recientemente la de HPV16 (Nadra *et. al.*, enviado). En todos los casos se conserva la topología de barril β -dimérico, pudiéndose calcular una desviación media (RMSD) de los C α entre 1 y 2 Å. Una comparación más detallada permite distinguir diferencias en la estructura cuaternaria entre las diferentes cepas, particularmente en la orientación relativa de los monómeros de E2c. Esto se evidencia al superponer las estructuras de un monómero de E2c y evaluar la orientación relativa del monómero no superpuesto. Como se muestra en la Fig 1.5, la desviación relativa de los monómeros permite separar en clases diferentes a HPV16 y HPV31 por un lado y a BPV1 por el otro. En este sentido, se hipotetizó que estas diferencias cuaternarias podrían influir en la unión a ADN, en tanto que la curvatura que debería sufrir el centro del palíndromo de ADN tendría que de ser diferente⁴⁸.

PLEGAMIENTO DE E2C

E2c es además modelo de estudios básicos de plegamiento proteico, en particular de plegamiento de dímeros. Los estudios iniciales mostraron que la dimerización de E2c esta acoplada al plegamiento de los monómeros, es decir, en solución sólo pueden encontrarse monómeros desplegados o dímeros plegados⁷³. Son dos cadenas polipeptídicas que conforman un sólo dominio de plegamiento independiente. Se describió también que este dímero es muy estable, con una constante de disociación en el orden subnanomolar, estabilizándose aún más a pH neutro. Dados estos antecedentes y las condiciones en las cuales trabajamos, nos referiremos aquí siempre a E2c como al dímero E2c.

La descripción del camino cinético de plegamiento de E2c identificó la formación de un intermediario monomérico, de rápida formación y desaparición. Se identificó también que este intermediario tiene características que lo identifican como no-nativo⁷². El efecto de la fuerza iónica y los ligandos reveló que el dímero E2c se estabiliza con el aumento de la fuerza iónica, y que esta estabilización es mayor aún si se utilizan sales de fosfato⁵⁵. Tanto la unión específica de ADN como la unión inespecífica de heparina estabilizan fuertemente el dímero frente al evento de desplegamiento/disociación inducido por urea⁵⁵.

Más recientemente, se describió el desplegamiento de E2c frente a la desnaturalización por presión hidrostática ²⁷. Se identificó que los monómeros desplegados por presión presentan estructura residual, y que éste estado es más compacto que el estado desplegado obtenido por desnaturalización con urea ²⁷. Los estudios estructurales por RMN sobre el estado desplegado en urea, muestran que éste también presenta estructura residual. Fue posible asignar que este estado es monomérico y que las zonas de la hélice-1 y la β 2 presentan estructuras no-nativas, tipo "vuelta" (*turn*) ⁷¹. Posteriormente se propuso que una vez unido a ADN, el dímero de E2c puede ser desplegado por presión hidrostática en monómeros que se mantienen unidos a ADN ³⁶.

Por otro lado, estudios de resonancia magnética nuclear permitieron asignar un alto grado de movilidad a las regiones de E2c involucradas en la unión a ADN. Los experimentos de intercambios protón/deuterio realizados sobre el homólogo de HPV31 muestran que tanto el *loop* β 2- β 3 como las hélices de reconocimiento presentan altos grados de movilidad local cuando el dímero no está unido a ADN ³⁴. Experimentos recientes de dinámica de relajación de protones amídicos, realizados con el homólogo de HPV16, muestran que esta cepa también presenta alta flexibilidad local en las zonas de unión a ADN, sugiriéndose además que el efecto podría estar controlado a nivel del N-cap de la hélice de reconocimiento (Nadra *et. al.*, enviado).

ANTECEDENTES DE UNIÓN E2c-ADN

Los primeros estudios de unión E2c-ADN permitieron la identificación de los "sitios específicos" de unión de E2. Como se mencionó, los HPV poseen 4 sitios E2 en el genoma, conservados tanto en secuencia como en posición relativa. Estos consisten en palíndromos imperfectos del tipo: ACCGnnnnCGGT, (la minúscula representa una base preferida, pero no requerida, y n cualquier base) ³⁵. Estudios posteriores destacan que, para las cepas de HPV, la secuencia consenso debería modificarse a: aACCGnnnnCGGT ¹⁰⁰. La región central, "nnnn", será referida como "separador", y cada región a los lados como "hemisito". Las estructuras cristalográficas muestran que los contactos específicos entre los aminoácidos de E2c y los motivos de las bases se dan sólo por los hemisitios (ver cap 4). Si bien no hay un consenso en la secuencia del "separador", las secuencias de los sitios de HPV tienden a estar enriquecidos en A:T. En 1990 Bendrosian y Bastia mostraron que HPV16-E2 une con mayor afinidad a los sitios enriquecidos en A:T ³. Por otro lado, Hines y col. proponen que HPV16-E2 posee una clara preferencia por separadores AA[AT]N, y que esta diferencia no está dada simplemente por el contenido de A:T, ya que un separador TTAA es unido con menor afinidad que uno AATT ³⁹. En el mismo trabajo se muestra que esto es una propiedad intrínseca a E2c de HPV16, ya que BPV1-E2 no presenta esta capacidad de discriminación (ver capítulo 2). En base a estos datos y a las estructuras resueltas, se propuso que la diferencia de afinidad de ambas cepas podría deberse a las restricciones conformacionales en la proteína y en el ADN, entendidas como

variantes de fenómenos de *lectura indirecta* de secuencia. En este sentido, se han cristalizado y resuelto estructuras de ADN de sitios E2 solos, en donde se destaca que las secuencias separadoras A:T pueden presentar grados de curvatura mayores que las ricas en C:G ⁴⁰. Recientemente se describió que la misma propensión a curvarse que presentan los separadores A:T se da también en solución ¹¹².

Las afinidades reportadas para la interacción E2c-ADN específico, varían entre 10^{-8} y 10^{-13} M ^{5, 39, 41, 87}. En general, estas mediciones se realizaron por ensayos de retardo de movilidad electroforética (EMSA o gel-shift), con diferentes condiciones de entorno, por lo que los valores no son comparables entre sí. Para un mismo ensayo se ha reportado la unión diferencial a los 4 sitios E2 del genoma de HPV16, proponiéndose un mecanismo de regulación de la transcripción por ocupación diferencial de los sitios E2 ⁸⁷. En el mismo trabajo se reportan algunos estudios cinéticos de la disociación de E2c de sitios específicos, destacándose una alta estabilidad de los complejos E2c-ADN específico, variable según el sitio específico del que se trate.

Como se tienen estructuras resueltas de E2c tanto en ausencia como en presencia de ADN, se hipotetizó sobre de los cambios estructurales asociados a la unión E2c-ADN, tanto para las cepas HPV18 y BPV1 ³⁹. Se propone que durante la unión no hay cambios sustanciales en la estructura del barril β -dimérico, ni en la orientación relativa de las hélices de reconocimiento, sino que los cambios mayoritarios se dan sobre la zona del *loop* conector β 2- β 3. Esta región no presenta densidad electrónica en los cristales de E2c libre, por lo que fue propuesta como altamente flexible, hipótesis confirmada por los experimentos de RMN mencionados. Para la cepa BPV1, la estructura del co-cristal con ADN muestra que la zona del *loop* está ordenada e involucrada en contactos salinos con los grupos fosfato del ADN, particularmente contactando la zona del "separador" de los hemisitios (Fig 1.4). Para el co-cristal de HPV18-E2c:ADN, la zona del *loop* no presenta densidad electrónica. Estos son algunos de los puntos que serán discutidos en mayor profundidad en los capítulos subsiguientes de la presente tesis.

MECANISMO DE UNIÓN AL EQUILIBRIO

"Well, there are two kinds of biologists, those who are looking to see if there is one thing that can be understood and those who keep saying it is very complicated and that nothing can be understood... you must study the simplest system you think has the properties you are interested in."

Cyrus Levinthal

MECANISMO DE UNIÓN AL EQUILIBRIO

En el marco biofísico planteado para disectar la asociación proteína-ADN, se requiere de un ensayo de medición de la actividad de unión a ADN que permita extraer parámetros fisicoquímicos y termodinámicos precisos. Para ello, el ensayo a utilizar debe cumplir con las premisas de los modelos a los cuales se lo va a ajustar: debe ser un método cuantitativo llevado a cabo en solución, con una mínima cantidad de componentes de entorno (sales, detergentes, glicerol, etc) y fundamentalmente, debe proveer las condiciones para que la unión de los ligantes sea reversible. En lo posible la medición debe ser no invasiva, es decir, permitir la cuantificación de las especies presentes (unido/libre) sin alterar el equilibrio mismo que se desea estudiar. Asimismo, las condiciones fisicoquímicas del entorno deben permanecer constantes durante el experimento, siendo preferible contar con la posibilidad de variarlas para los diferentes ensayos. Desde ya que para realizar todas estas mediciones se debe además contar con la proteína y el ADN en un alto grado de pureza, cuantificados con precisión.

Los métodos comunmente utilizados para la medición de la actividad de unión a ADN requieren de la separación física entre los reactantes libres y los complejos proteína-ADN. Uno de los más utilizados es el ensayo de retardo electroforético, EMSA o gel-shift. Si bien este ensayo resulta muy útil para una estimación cualitativa de la actividad de unión, incluso en extractos no purificados, durante la corrida electroforética las condiciones de entorno del gel no permanecen constantes ni controlables, pudiendo introducir importantes errores en la cuantificación de la interacción ³⁸. Otro de los métodos desarrollados es el ensayo de retención de complejos proteína-ADN en filtros de nitrocelulosa. En éste es posible cuantificar la unión en solución por medio de una marca radioactiva en el ADN, pero resulta poco atractivo dada la gran cantidad de material que se debe utilizar y la poca resolución temporal que brinda.

El método elegido para seguir la actividad de unión, y al que se referirá de aquí en adelante, es un método fluorimétrico basado en el cambio de las propiedades fluorescentes de uno o más fluoróforos. Cuando se requirió que la marca se encuentre en el ADN, se unió una molécula de isotiocianato de fluoresceína (FITC) via un "brazo" de 6 carbonos al extremo 5' de los oligonucleótidos (ver capítulo 7). Cuando se requirió la marca en la proteína, se utilizaron los Trp de la misma, es decir la fluorescencia intrínseca de E2c.

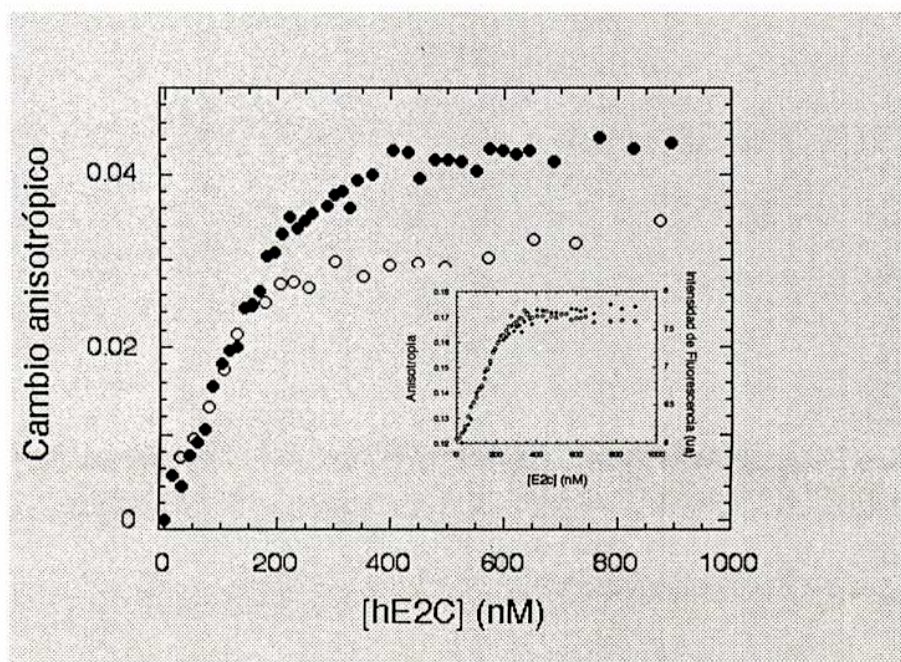
La puesta a punto del ensayo de unión permitió encontrar condiciones en las cuales la unión es reversible, los reactantes libres y el complejo proteína-ADN son solubles, y la unión esta asociada a un cambio fluorescente. Las mínimas condiciones de entorno encontradas requieren el uso de 25 mM BisTris pH7.0, 1mM DTT, y 200 mM NaCl, mínima cantidad de fuerza iónica requerida para evitar la formación de agregados insolubles en estas condiciones.

ESTEQUIOMETRÍA DE UNIÓN - EVIDENCIA ESPECTROSCÓPICA

El estudio detallado de la interacción requiere conocer la estequiometría de reacción. El oligonucleótido de referencia utilizado es el "site35" (s35): un oligonucleótido de doble cadena, de 18pb, conteniendo la secuencia del "sitio 35" de unión de E2, presente en el genoma de HPV-16 (ver capítulo 7). El FI-s35 está derivatizado con FITC en el extremo 5' de una de las hebras. La Fig 2.1 muestra los cambios de las propiedades fluorescentes del FI-s35 en una titulación con hE2c a una concentración constante de 200nM ADN. La anisotropía de la fluorescencia del motivo FITC aumenta a medida que se agrega hE2c, evidenciando la formación del complejo proteína-ADN. Este aumento es atribuible al mayor volumen hidrodinámico del complejo, y por tanto del coeficiente de correlación rotacional de la macromolécula que lleva la marca fluorescente, el ADN. El aumento de la anisotropía llega a una concentración de proteína de 200nM (relación molar 1:1) y satura a 400nM (2:1 proteína-ADN). El cambio anisotrópico total es de 0.045 unidades, de las cuales aproximadamente la mitad cambian a la relación equimolar. Hay un pequeño cambio de intensidad de la fluorescencia de FITC, que va en paralelo con el cambio en la anisotropía de la fluorescencia (Fig 2.1 inserto). Dado que el cambio de intensidad es menor al 15% de la intensidad original, no fue necesario corregir los valores anisotrópicos por el cambio del rendimiento cuántico de la fluoresceína ⁵².

Realizando la titulación en las mismas condiciones pero a pH 5.5, la señal de unión aumenta linealmente hasta 200nM, saturando con el agregado de sólo un equivalente de hE2c. La amplitud del cambio anisotrópico es la misma que la observada para la unión de un equiva-

Fig 2.1: Estequiometría de unión hE2c:s35. FI-s35 fué incubado con concentraciones crecientes de hE2c en las condiciones estándar a pH 7.0 (puntos negros) o pH 5.5 (puntos blancos). Inserto: cambios de intensidad y anisotropía de la fluorescencia de la fluoresceína a pH 7.0



lente a pH 7.0. Tomados en conjunto, estos datos permiten estimar que la preparación recombinante de hE2c tiene una actividad de unión cercana al 100% (ver capítulo 7). Por otro lado, los datos sugieren que la estequiometría de unión es de 1 dímero de hE2c por sitio de unión cuando el ensayo se realiza a pH5.5 y de dos dímeros de hE2c por sitio, si el ensayo es a pH 7.0.

La estequiometría encontrada a pH 7.0 es inesperada dados los modelos presentados para la unión entre hE2c y ADN descritos por otros grupos de trabajo, en particular las estructuras cristalográficas ³⁵. Buscando confirmar o descartar la estequiometría encontrada, realizamos una titulación en las mismas condiciones pero en el sentido inverso, utilizando a hE2c como ligando fijo y agregando s35. La señal seguida en este caso es el aumento de la intensidad de fluorescencia de los Trp de hE2c, evento ya caracterizado como asociado a la unión de ADN ⁵⁵. En el inserto de la figura 2.2 se muestran los resultados de la titulación, donde se observa un aumento lineal de la intensidad hasta una relación molar 0,5:1 (s35:hE2c), y un aumento lineal, de menor pendiente, luego de la transición. Esto sugiere, de forma independiente, que la estequiometría de reacción es de 2 moléculas diméricas de hE2c por sitio de 18pb, descartando la posibilidad de que la detección de 2 eventos de unión se deba al motivo FITC ligado al oligonucleótido.

El dominio C-terminal de E2 de HPV16 (hE2c) tiene 3 Trp por monómero, dos en el centro del barril- β (W345, W347) y uno en la hélice-2 (W339), expuesto al solvente, y del lado opuesto del barril- β respecto al sitio de unión de ADN (Fig 1.4). Para diferenciar las contribuciones de unos u otros cromóforos intrínsecos, construimos una mutación sitio-dirigida sobre W339, reemplazándolo por Phe (W339F). La proteína fue expresada y purificada y se mostró idéntica a la salvaje en cuanto a estabilidad, estructura secundaria y unión a ADN (no mostrado). Esta mutante también presenta un cambio de intensidad de fluorescencia de Trp al unir ADN, lo cual indica que existe algún tipo de rearrreglo en el barril- β cuando hE2c une ADN. La isoterma de unión de W339F y ADN a alta

Figura 2.2: Estequiometría de unión hE2c:s35.
200nM hE2c fue incubado con concentraciones crecientes de ADN s35 en las condiciones de referencia a pH 7.0. Se registró el cambio en la intensidad de fluorescencia de los Trp de hE2c W339F a 330nm. Inserto: mismos cambios para hE2c salvaje

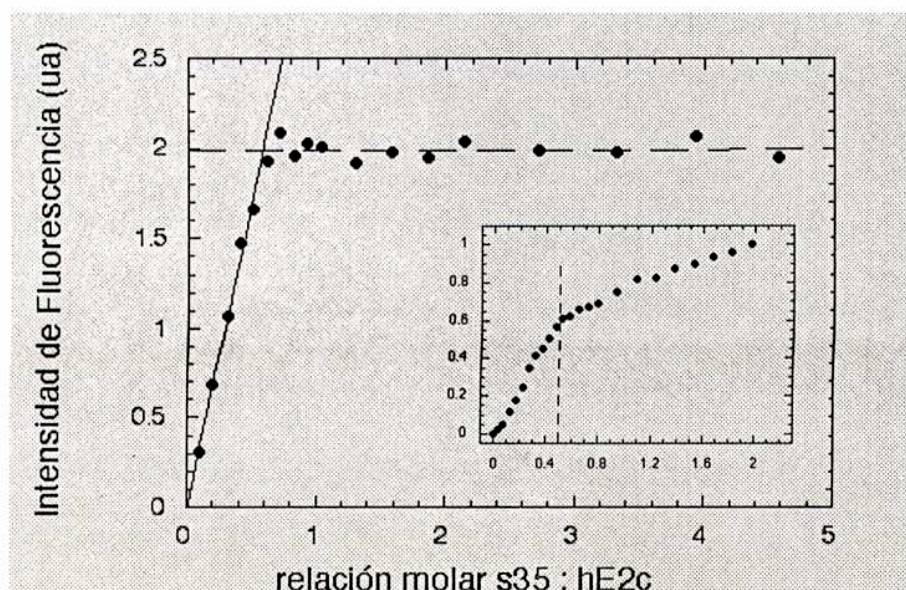
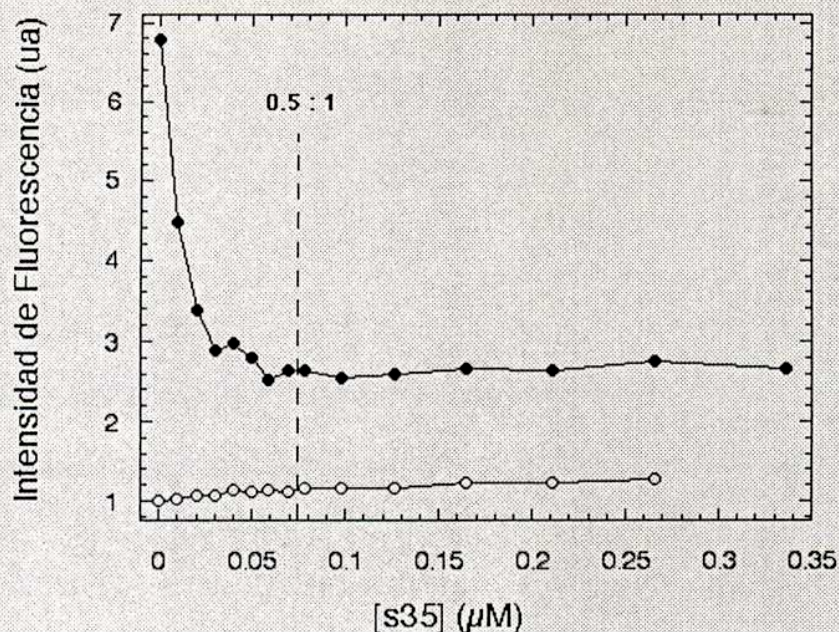


Fig 2.3: Estequiometría de unión hE2c:s35 seguida por desplazamiento de bis-ANS.
600nM bis-ANS fué incubado con concentraciones crecientes de s35 en presencia (puntos negros) o ausencia (puntos blancos) de 150nM hE2c. Se grafica la intensidad de fluorescencia de bis-ANS integrada entre 400 y 600 nm.



concentración se muestra en la Fig 2.2. Como es evidente, la estequiometría de unión es también de 2:1 (hE2c:s35), y no se observa ningún cambio fluorescente posterior que, para la proteína salvaje es atribuible al W339.

Estudios previos mostraron que hE2c une dos moléculas de bis-ANS por monómero, una de las cuales es desplazada por la unión de ADN a hE2c⁵⁵. Una forma independiente de confirmar la estequiometría de reacción es medir la estequiometría del desplazamiento de dicha sonda al unirse hE2c y s35. Se realizó una titulación en las mismas condiciones anteriores, en presencia de bis-ANS a concentraciones de saturación de los sitios en hE2c. La Fig 2.3 muestra que todo el bis-ANS desplazable por ADN es expulsado a una relación 0,5:1 de ADN:proteína, confirmando la estequiometría encontrada anteriormente, con una sonda fluorescente diferente. Un experimento control muestra que la intensidad de fluorescencia del bis-ANS no varía con el agregado de s35 (Fig 2.3).

ESTEQUIOMETRÍA DE UNIÓN - EVIDENCIA NO ESPECTROSCÓPICA

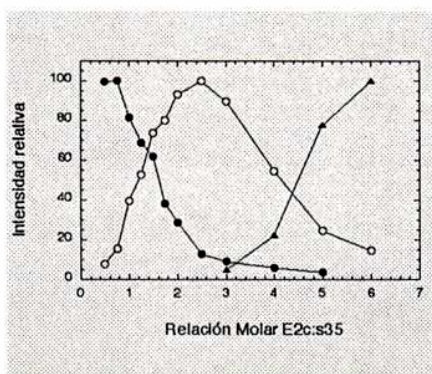
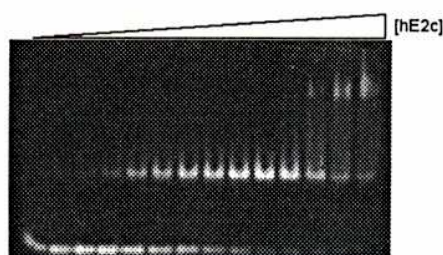


Figura 2.4: Estequiometría de unión hE2c : FI-35 analizada por electroforesis en gel: 2 μ M FI-s35 fue incubado con concentraciones crecientes de hE2c y las muestras se sembraron en un gel de poliacrilamida al 6% en condiciones nativas.
A) visualización directa de la fluorescencia de FI-s35.
B) cuantificación de las bandas: ADN libre (circuitos negros), primera banda de retardo (circuitos blancos), segunda banda de retardo (triángulos negros).

La inesperada estequiometría de unión descrita para los ensayos a pH 7.0 (dos moléculas diméricas de hE2c por sitio E2 canónico), nos llevó a buscar evidencia no-espectroscópica de la formación de los diferentes complejos. Se ensayó discriminar la formación de las especies por filtración molecular pero, en las condiciones de unión establecidas, la proteína interactúa con la matriz y es retenida durante la filtración. Para desestabilizar esa interacción y eluir los complejos fue necesario aumentar la fuerza iónica, pero eso mismo influye muy fuertemente en la unión proteína:ADN (ver mas abajo). El método elegido para separar las especies fue la separación por electroforesis en geles nativos, EMSA. FI-s35 fue incubado con concentraciones crecientes de hE2c y los complejos resueltos en un gel de poliacrilamida en condiciones de baja fuerza iónica ² (Fig 2.4A). En el primer carril se distingue la banda correspondiente al ADN libre, que pierde intensidad a medida que aumenta la concentración de hE2c, dando lugar a la aparición de una primera banda de retardo. Si la concentración de hE2c se aumenta aún más, se observa que la primera banda desaparece, dando lugar a la formación de otra banda de mayor retardo, que atribuimos a la formación de un complejo 2:1 hE2c:ADN. La aparición de fluorescencia difusa entre ambas bandas es evidencia de que la estabilidad del complejo 2:1 es menor que la del 1:1 en las condiciones en las cuales se realizó la electroforesis.

La cuantificación de las bandas observadas muestra la existencia de un equilibrio entre las poblaciones sugeridas: el complejo 1:1 se incrementa a medida que desaparece el ADN libre y este desaparece a medida que se acumula la especie 2:1, sugiriendo que la segunda especie se forma a expensas de la primera (Fig 2.4B). La banda correspondiente al complejo 1:1 satura alrededor de 2,5 μ M, permitiendo estimar una constante de disociación (K_D) aparente cercana a 1 μ M. Dadas las mediciones posteriores de la afinidad de hE2c por el FI-35 realizadas en solución, esto indica que las condiciones electroforéticas utilizadas son disociantes, probablemente por la falta de "aditivos" (cationes, detergentes, etc) a la mezcla de reacción ³⁸. Cabe mencionar que el mismo gel fue luego teñido con Ag para evidenciar la presencia de proteína en los supuestos complejos. Ambas bandas de retardo fueron teñidas, a diferencia de la del FI-35 libre (no mostrado).

MEDICIÓN DE LA ENERGÍA LIBRE DE
UNIÓN
INTERACCIÓN hE2c - ADN ESPECÍFICO

Las titulaciones presentadas en la sección anterior muestran la existencia de dos eventos de unión de hE2c sobre el s35. Si la titulación se realiza a 2nM FI-s35 constante, el cambio anisotrópico total es de 0.02 unidades, cerca de la mitad del cambio total detectado a 200nM. Esto nos sugirió que, trabajando a concentraciones bajas, es posible separar los eventos de unión manteniendo las condiciones de solubilidad y reversibilidad antes descritas. La baja concentración de la marca fluorescente impide el uso de polarizadores para las mediciones anisotrópicas, dada la menor intensidad lumínica y la excitación diferencial de una baja proporción de los fluoróforos totales.

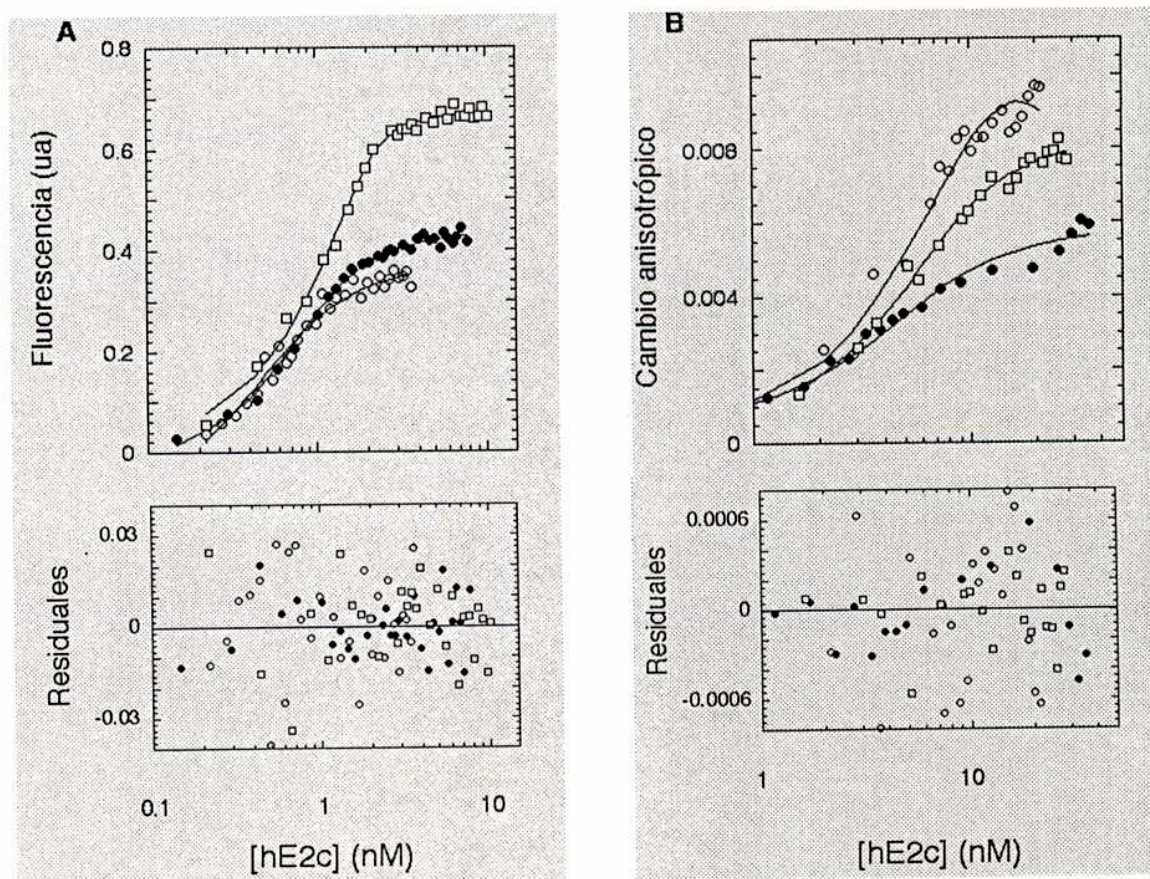


Figura 2.5: Determinación de la energía de interacción entre hE2c y FI-s35. Los experimentos de unión se llevaron a cabo agregando hE2c a una solución de FI-s35 en las condiciones estándar. A) Unión de alta afinidad: 0.5 nM (círculos blancos), 1.0 nM (círculos negros) o 2.0 nM FI-s35 (cuadrados blancos). El panel de abajo muestra la distribución de los residuales al ajuste. Las K_D extraídas son 0.19 ± 0.04 ; 0.17 ± 0.02 ; 0.18 ± 0.02 nM respectivamente. B) Unión de "baja" afinidad: 5nM (círculos negros), 7nM (cuadrados blancos) o 10 nM FI-s35 (círculos blancos). El panel de abajo muestra la distribución de los residuales al ajuste. Las K_D extraídas son 1.6 ± 0.4 ; 1.4 ± 0.2 ; 2.1 ± 0.3 nM respectivamente.

Las mediciones se realizaron entonces siguiendo el cambio asociado de la intensidad de fluorescencia de FITC. Las isothermas de unión a concentraciones fijas de ADN de 0.5 , 1 y 2 nM se muestran en la figura 2.5A. Los datos fueron ajustados a un modelo de unión simple, asumiendo una estequiometría de 1:1 (ecuación 7.1, capítulo 7). Las K_D 's extraídas de las curvas son de 0.19 ± 0.04 ; 0.17 ± 0.02 ; 0.18 ± 0.02 nM respectivamente. La coincidencia en el valor de K_D de las mediciones independientes realizadas a diferentes concentraciones de ligando fijo, junto a la distribución aleatoria de los residuales, validan el modelo de unión utilizado y los valores de K_D obtenidos (Fig 2.5).

Para la medición de la energía libre del segundo evento de unión, la estrategia elegida fue utilizar como ligando fijo un complejo 1:1 hE2c-Fis35. Las isothermas de unión al agregar hE2c y el posterior ajuste al mismo modelo de unión, nos permitió extraer una K_D media de 1.8 ± 0.5 nM de tres experimentos independientes (Fig 2.5B). Interpretamos que el evento fluorescente detectado es el asociado a la unión de un dímero de hE2c a un complejo 1:1 hE2c-Fis35 preformado. En esta etapa, no es posible discernir si la unión de la molécula adicional de hE2c se da sobre el ADN, la proteína unida o ambos.

INTERACCIÓN hE2c - ADN INESPECÍFICO

La funcionalidad de las proteínas que unen ADN específicamente depende de la unión diferencial a un sitio concreto frente a un gran exceso de potenciales sitios de unión no específicos. Es por esto que cualquier interpretación de una medida de unión a un sitio específico debe ser tomada en términos relativos a la unión de la misma a sitios no relacionados. Medimos la afinidad de hE2c por oligonucleótidos de doble cadena de la misma longitud que el referente s35, pero con diferentes secuencias de bases. Las secuencias utilizadas corresponden a un sitio de unión de la proteína EBNA1 (Epstein Barr Nuclear Antigen 1), proteína estructuralmente relacionada con hE2c, pero sin relación evolutiva cercana ⁷ (ver capítulo 1). El otro oligonucleótido utilizado, iset35, fue generado al azar manteniendo la misma composición de bases y longitud que s35. Las interacciones fueron medidas por el mismo método y en las mismas condiciones de solución que las descritas para el FI-s35. Para los casos de uniones inespecíficas, la estequiometría de unión observada es de 1:1, si bien no se puede descartar la formación de complejos 2:1 a mayores concentraciones que las ensayadas. Se realizaron curvas de unión a diferentes concentraciones de ligando fijo y las constantes calculadas se presentan en la tabla 2.1. La capacidad de discriminación de secuencia, definida como la relación K_D inespecífico/específico es de 10000 para el oligonucleótido EBNA y 4000 para iset35, en el orden de lo reportado para E2c por otros grupos ³⁵ y el orden esperado para interacciones específicas proteína-ADN en general ¹⁰⁵.

	site 35 K_D (nM) (cognato HPV-16)	site bov K_D (nM) (cognato BPV-1)	ebna K_D (nM) (noespecífico)	Índice de discriminación ^b	Índice de especificidad ^c
HPV-16 E2C ^a	1: 0.2 ± 0.02 2: 2 ± 0.3	1: 30 ± 15 2: 3000 ± 1200	2000 ± 60	150	10000
BPV-1 E2C	700 ± 70	200 ± 23	11000 ± 1600	3.5	55

Tabla 2.1: Afinidades y Discriminación de secuencia de la Interacción hE2c - ADN y bE2c - ADN.
ojo, meter isel35.

a: para HPV-16E2c se muestran las constantes de disociación de los dos "modos".

b: calculado como la constante de disociación por oligonucleótido no-cognato / cognato.

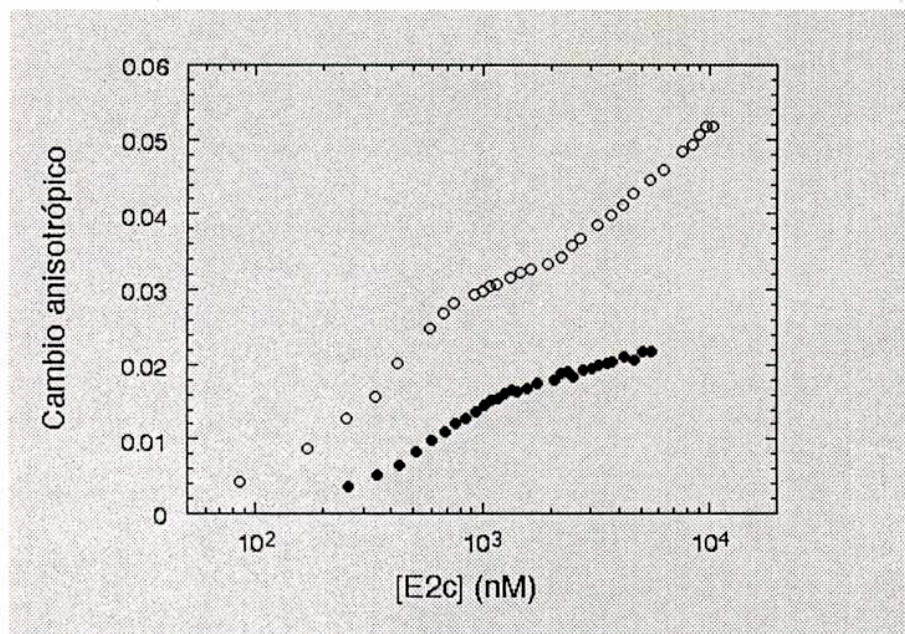
c: Calculado como la constante de disociación por oligonucleótido inespecífico / cognato.

COMPARACIÓN DE LA ACTIVIDAD DE UNIÓN A ADN DE E2C PARA LAS CEPAS HPV16 Y BPV1

Gran parte de los estudios donde la proteína E2 es utilizada como modelo han sido realizados con el homólogo del virus del papiloma bovino 1 (bE2). Como se detalla en el capítulo 1, esta proteína esta estrechamente relacionada con hE2c en cuanto a su distribución de dominios y topología general de plegamiento. Es también la proteína con la cual se han realizado las mayores contribuciones respecto al mecanismo de regulación de la transcripción de los *papilomavirus*, constituyéndose en el modelo genético más estudiado ⁶⁸. Es por esto que iniciamos un análisis comparativo de hE2c y bE2c en cuanto a sus características de unión a ADN en solución, utilizando la metodología desarrollada.

Con este fin clonamos, expresamos y purificamos el dominio bE2c utilizado para estudios cristalográficos ³⁴, como se detalla en el capítulo 7. Utilizamos como oligonucleótido blanco el sBPV, también de 18pb, que contiene un sitio de unión E2. La secuencia es la misma que la del oligonucleótido utilizado para los estudios cristalográficos, difiriendo de s35 en el "separador" de los hemisitios del sitio canónico E2. Las condiciones de unión puestas a punto para hE2c (pH=7.0, 0,2M NaCl), son suficientes para mantener la solubilidad de los reactantes y productos, y da lugar a una unión reversible (no mostrado).

Figura 2.6: Interacción de hE2c y bE2c con el oligonucleótido sBPV. 500nM FI-sBPV fue incubado en las condiciones estándar con concentraciones crecientes de bE2c (círculos negros) o hE2c (círculos blancos). Se muestra el cambio de la anisotropía de la fluorescencia.



Las K_D de la interacción de ambos dominios (hE2c, bE2c) por sBPV se obtuvieron en condiciones de disociación, y los valores extraídos se muestran en la tabla 2.1. Como se mencionó, hE2c presenta dos eventos de unión a sBPV, siendo la relación de las afinidades entre ambos de 100 veces, 10 veces mayor que la relación entre los valores análogos para s35. Complementando la comparación, se muestran los valores obtenidos para la interacción la proteína bE2c con los sitios s35 y EBNA. La relación entre las constantes obtenidas da idea de la capacidad de discriminación entre sitios cognatos, no-cognatos e inespecíficos de la misma longitud. Considerando el evento de mayor afinidad, hE2c discrimina 150 veces por sitios cognatos, y 10000 veces por sitios específicos. Por otro lado, bE2c presenta una preferencia 3.5 veces mayor por el sitio cognato y sólo 55 veces por un sitio específico.

CAMBIOS CONFORMACIONALES SOBRE EL ADN INDUCIDOS POR E2C

Las diferencias más notables entre los sitios E2 de s35 y sBPV residen en la secuencia separadora de los hemisitios. Coincidiendo con lo presentado aquí, se ha descrito que hE2c une con más afinidad sitios con separadores ricos en A:T^{5, 39}. Dado que las bases contactadas son las mismas (los hemisitios), se hipotetizó que las diferencias de afinidad se podrían deber a la flexibilidad diferencial que presentarían estos sitios E2. Según las estructuras resueltas para los complejos, el ADN del sitio E2 se curva centrado con la secuencia "separadora", no contactada por la proteína (Fig 1.4). Si, como se propuso, hE2c y bE2c presentan diferencias en la orientación relativa de las hélices de reconocimiento, menos energía será requerida para unir sitios de ADN flexibles (con secuencias A:T) como el s35. Una predicción de este modelo es que las estructuras del ADN unido a hE2c o bE2c, deben presentar diferencias estructurales relacionadas con parámetros de curvatura y/o apilamiento de bases.

El dicroísmo circular en el UV-cercano es una herramienta útil para analizar cambios en la conformación de oligonucleótidos de ADN en solución³⁰. En las mismas condiciones que las utilizadas para la medición de la energía de unión, pero a concentraciones muy por encima de las K_D , la conformación del s35 cambia al unir un equivalente de hE2c (Fig 2.7A).

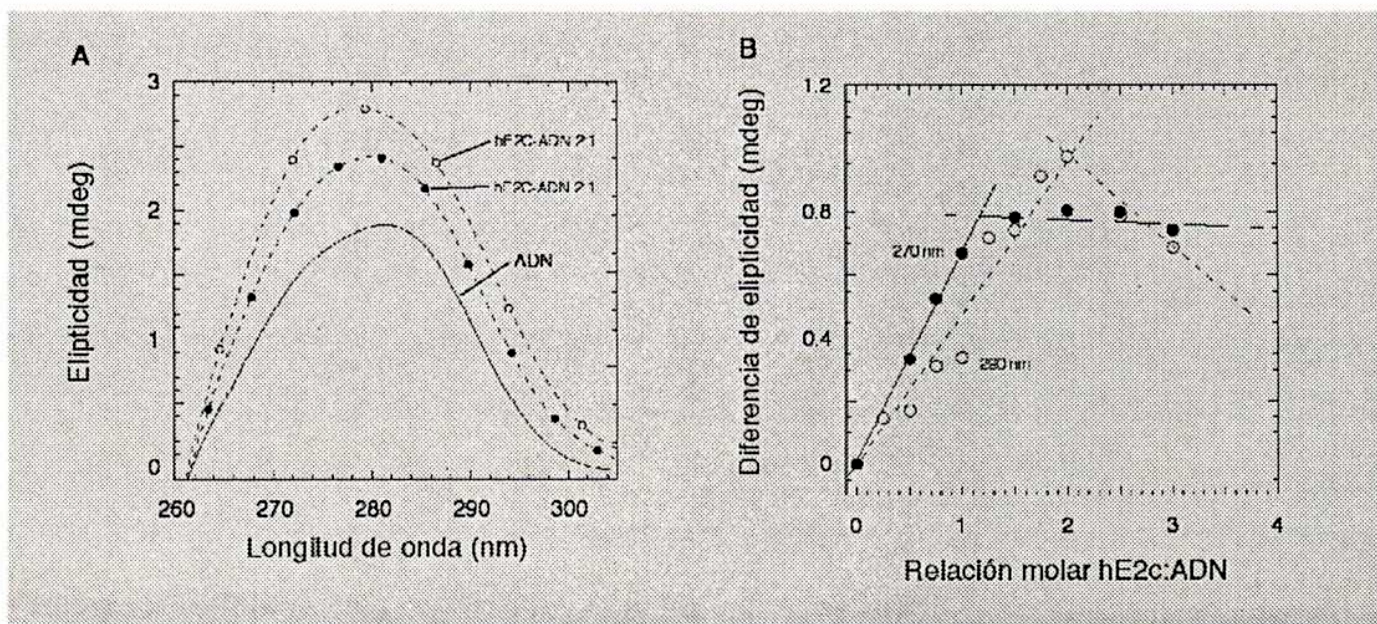
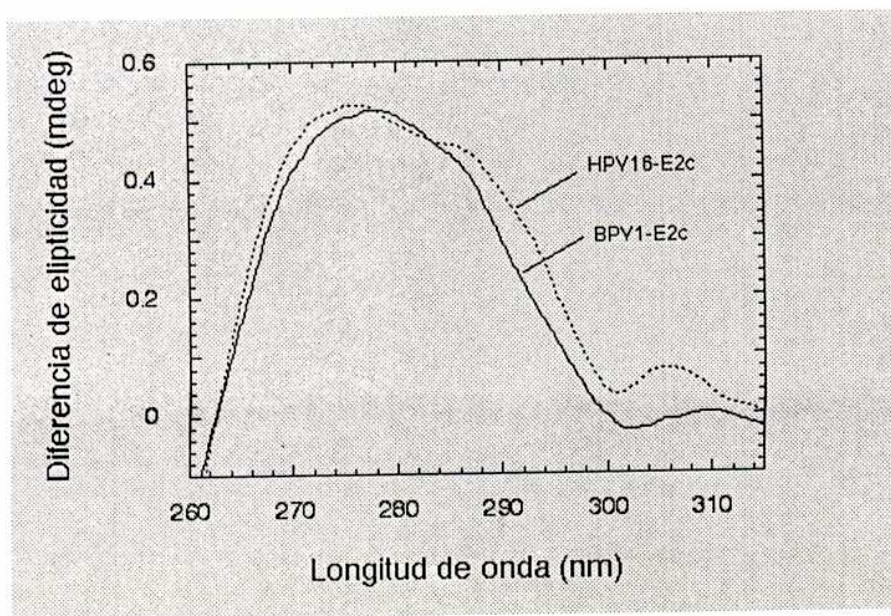


Figura 2.7: cambios conformacionales sobre ADN s35 al unir hE2c. A) 1 μ M s35 fue incubado en las condiciones estándar y se registró el espectro de dicroísmo circular. s35, línea sólida; con 1 μ M hE2c; puntos negros; 2 μ M hE2c, puntos blancos. B) titulación de 1 μ M s35 con hE2c siguiendo la elipticidad a 270nm (puntos negros) o 292 nm (puntos blancos).

Dada la baja señal de elipticidad molar que presenta hE2c en esta región, el cambio es mayoritariamente atribuible a rearrreglos conformacionales en el ADN. El incremento de la elipticidad entre 270 y 290nm puede ser asignado a un desenrollamiento parcial de la doble hélice y/o a un desapilamiento de pares de bases ³⁰, en plena concordancia con las evidencias cristalográficas ³⁹. El agregado de otro equivalente molar de hE2c muestra que la elipticidad continúa incrementándose, siendo esto evidencia independiente de la existencia de un evento adicional de unión. El seguimiento detallado de los cambios dicróicos muestra que la elipticidad a 270nm satura a un estequiometría 1:1 hE2c:s35, mientras que a 292nm lo hace a 2:1 (Fig 2.7B). El hecho que las bandas de dicróismo se comporten de manera diferencial al unirse 1 o 2 dímeros de hE2c, muestra que el oligonucleótido sufre rearrreglos conformacionales distintos al unir uno o dos dominios hE2c, sugiriendo que los eventos detectados podrían tener significancia funcional.

En la sección anterior mostramos que bE2c une ADN con menor afinidad que hE2c, incluso por sitios cognatos de BPV1 (tabla 2.1). Un estudio presentado por el grupo de Hegde y col.³⁹ muestra que el dominio hE2c tiene mayor afinidad por sitios que presentan una mayor "flexibilidad". Particularmente, proponen además que hE2c es más dependiente que bE2c a modificaciones en la secuencia "separadora" de los hemisitos del sitio E2. Basados en evidencia cristalográfica³⁶, sugirieron que hE2c y bE2c deben deformar el ADN diferencialmente para "acomodar" las hélices de reconocimiento en los zurcos mayores. En este sentido, y buscando evidencia espectroscópica en solución, medimos los cambios conformacionales sobre el s35 al unirse hE2c o bE2c (Fig 2.8).

Figura 2.8: cambios conformacionales inducidos por hE2c o bE2c sobre s35. 1µM s35 fue incubado en las condiciones estándar con el agregado de 1µM hE2c (línea punteada) o 1µM bE2c (línea sólida). Se muestran los espectros diferencia.

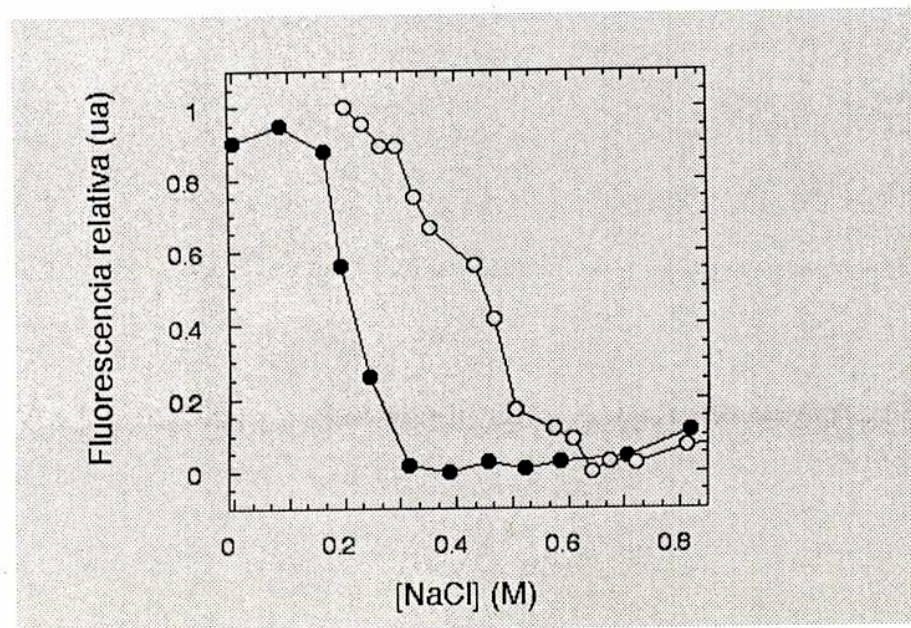


EFFECTO DE LA FUERZA IÓNICA EN LA INTERACCIÓN

Los espectros de los cambios diferenciales de DC de bE2c muestran un corrimiento al rojo de la banda de 275nm, y un disminución de la intensidad de la banda de 292nm, respecto al complejo hE2c:s35. Se diferencia además una banda a 305nm que no esta presente para el complejo bE2c-s35. Teniendo en cuenta que el ADN es el mismo, los resultados muestran que ambas proteínas distorsionan el mismo sitio de reconocimiento de forma diferente, si bien no es posible extraer información estructural detallada de estos espectros.

A pesar de la cercana homología y apariencia estructural, bE2c y hE2c presentan sutiles diferencias estructurales. Una de ellas esta asociada al potencial electrostático generado por la exposición superficial de cargas en la región de unión de ADN. En las mismas condiciones, hE2c esta menos (positivamente) cargada que bE2c³⁶. Dadas los tipos de interacciones que se dan entre proteínas y ADN, esta generalmente aceptado que cuanto mayor es el componente electrostático, más debil y menos especifica es la unión. Los resultados de unión y discriminación presentados, junto con la evidencia cristalográfica, sugieren que la contribución electrostática a la unión de bE2c sería mayor que la de hE2c. Esta hipótesis fué ensayada midiendo la disociación del complejo E2c-ADN ante el incremento de la fuerza iónica de la solución (Fig 2.9).

Figura 2.9: Sensibilidad de la unión E2c-ADN a la fuerza iónica. 500 nM de complejos 1:1 FI-s35 hE2c (puntos blancos) o FI-s35-bE2c (puntos negros) fueron incubados con concentraciones crecientes de NaCl. Se muestran los cambios de intensidad de fluorescencia.



Se observa que la unión de bE2c a s35 es más dependiente de la fuerza iónica que la de hE2c, confirmándose que la unión bE2c-ADN tiene un mayor componente electrostático. El punto medio de la transición de hE2c-s35 es de 0.45 M NaCl, mientras que para bE2c, 0.2 M NaCl es suficiente para disociar la mitad de las moléculas del complejo con ADN.

MODELO DE UNIÓN AL EQUILIBRIO

Dados los resultados descriptos, planteamos el siguiente modelo de unión al equilibrio entre hE2c y ADN, en las mínimas condiciones de entorno que permiten la solubilidad de los componentes y la reversibilidad de la interacción (Fig 2.10). Trabajos anteriores muestran que la constante de disociación del dímero de E2c en los monómeros componentes esta en el orden picomolar^{72, 73}. El dímero se une inespecíficamente a ADN con una constante del orden micromolar, y específicamente con constantes del orden subnanomolar; pudiendo estimarse un factor de especificidad de 1×10^4 . Es posible que tanto las constantes como el índice de especificidad cambien al cambiar las condiciones de reacción, por tanto no es posible asignar directamente estos valores a los que puedan ocurrir *in vivo*. Si la reacción se realiza a pH 7.0, es posible distinguir un segundo evento de unión al complejo hE2c:s35. Este tiene una constante de disociación del orden nanomolar, demasiado afín como para ser considerada como no específica. Durante ambos eventos se detectaron cambios en la estructura terciaria del ADN, compatibles con los resultados cristalográficos obtenidos para la unión con ADN de los homólogos de hE2c. Como ambos eventos presentan cambios en las señales de DC de ADN, podemos suponer en esta instancia que la unión del segundo dímero involucra contactos directos sobre el ADN, si bien no es posible asignar unívocamente el sitio de unión.

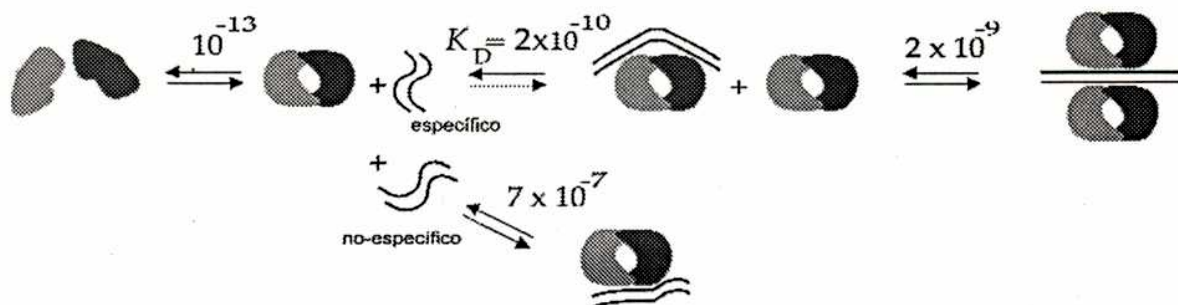


Figura 2.10: Modelo de unión hE2c:ADN a pH=7.0. Los monómeros desplegados se asocian con una constante subpicomolar, formando dímeros capaces de unir ADN. Se muestran las constantes de unión con ADN específico y no específico. Las bases estructurales del reconocimiento 2:1 restan ser definidas, por lo tanto el dibujo es puramente esquemático

MECANISMO CINÉTICO DE UNIÓN *E2c* - ADN

"Habría que sentir en velocidad"

H. M. Andrade

La disección del mecanismo cinético de interacción entre macromoléculas provee información fundamental sobre los eventos involucrados en el reconocimiento, así como de una relación temporal entre éstos. La inclusión del dominio de tiempo da cuenta de los intermediarios involucrados y de las rutas de sucesión de estados, permitiendo a su vez la caracterización de los "conjuntos de estados de transición" (TSE, por *transition state ensemble*) por los que la reacción transcurre. El uso de herramientas termodinámicas permite la integración de estos datos mecanísticos en términos estructurales, relacionando las estructuras proteicas con su funcionalidad bioquímica, y proveyendo una interpretación molecular detallada del mecanismo de unión y reconocimiento entre macromoléculas.

COLISIÓN INICIAL CON ADN ESPECÍFICO

Haciendo uso del ensayo de unión en solución desarrollado (capítulo 2), realizamos experimentos cinéticos de unión en condiciones de pseudo-primer orden respecto a E2c de HPV16, a la que nos referiremos de aquí en adelante como E2c. Utilizando un aparato de flujo detenido (stopped-flow), mezclamos Fl-s35 y E2c registrando el cambio de la intensidad de fluorescencia asociado a la formación del complejo proteína:ADN. La Fig 3.1A muestra el aumento de la intensidad de fluorescencia del motivo FITC en función del tiempo de reacción luego del mezclado. Los residuales del ajuste de los datos a decaimientos exponenciales simples evidencian que el proceso puede ser descrito como un decaimiento biexponencial (Fig 3.1A, inserto), vale decir, se corresponde con dos *fases* cinéticas. Como muestra la figura 3.1B, ambos eventos fluorescentes son dependientes de la concentración de E2c, indicando que en ambos casos se trata de procesos cinéticos de segundo orden, asociados a reacciones bimoleculares. Las constantes cinéticas de segundo orden (k_{on}) calculadas para ambos eventos de colisión son de $0.6 \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$ y $1.4 \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$. Es de destacar que la magnitud de estos valores es comparable a los calculados para el límite físico impuesto por la difusión de macromoléculas de tamaños comparables a E2c y s35^{6, 111}. El hecho de que las colisiones sean tan rápidas que la velocidad límite esté dada por la difusión, puede probarse evaluando la variación de las constantes cinéticas al aumentar la viscosidad de la solución. Como se muestra en la figura 3.1C, las velocidades disminuyen al aumentar la viscosidad. El hecho de que la pendiente de este gráfico tenga un valor cercano a la unidad refleja que se trata de colisiones bimoleculares limitadas por la difusión²⁵. El efecto causado sobre las velocidades de colisión es el mismo, tanto si la viscosidad se controla por el agregado de sacarosa (Fig 3.1C) o de glicerol a la solución (no mostrado).

El valor extrapolado para la disociación (k_{off}) del evento más rápido es de 6 s^{-1} , lo que permite estimar la constante de disociación al equilibrio ($K_D = k_{off}/k_{on}$) de 5nM para el "complejo de encuentro" formado luego de la colisión. La disociación del evento más lento (k_{-1}) parece aún menor, si bien no fue posible determinarla en estas condiciones.

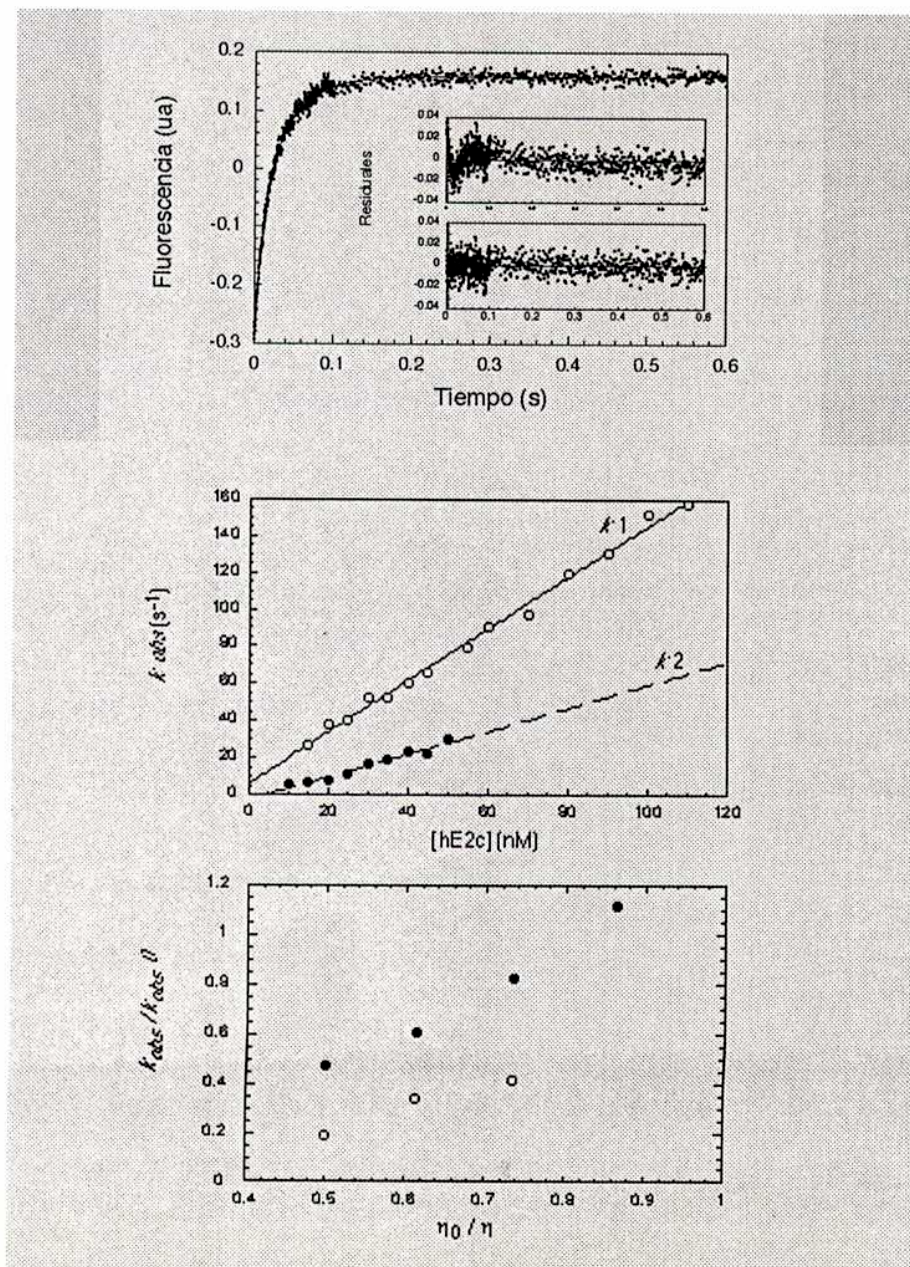


Figura 3.1: Asociación cinética de E2c y FI-s35.

A) curva de asociación cinética típica entre 5nM FI-s35 y 40nM E2c. La línea es el mejor ajuste de los datos a decaimientos exponenciales. Los residuales a ajustes de 1 o 2 exponenciales se muestran como inserto.

B) curva de pseudoprimer orden de las constantes extraídas de ambas fases a 5nM FI-s35, k_{on1} (círculos negros) y k_{on2} (círculos blancos). Por arriba de 60nM [E2c], los datos ajustan a 1 exponencial. C) dependencia de las constantes observadas con la viscosidad, k_{obs1} (puntos negros), k_{obs2} (puntos blancos).

De los dos eventos de asociación observados, la amplitud relativa del evento más lento (k_{on2}) disminuye a medida que la concentración de E2c aumenta (Fig 3.2A). A las concentraciones más altas de E2c ensayadas, los datos ajustan a un proceso exponencial simple, indicando que sólo una de las colisiones bimoleculares ocurre a altas concentraciones relativas. Una forma de explicar este comportamiento es plantear que esta ocurriendo una competencia entre los reactantes, de modo que a medida que aumenta la concentración, el reactivo que permanece fijo (ADN), se hace limitante. De esta forma, a altas concentraciones solo tendría lugar el proceso más rápido. Una alternativa es que la macromolécula variable (E2c) cambie al aumentar la concentración. No hay evidencias de que E2c forme multímeros en solución, incluso cuando se la estudia a concentraciones milimolares en experimentos de RMN. Por otro lado, en las condiciones en las cuales realizamos el ensayo, la constante de disociación del dímero de E2c está en el orden picomolar, permitiéndonos afirmar con confianza que la población de proteína es >99% dimérica⁷³. Teniendo en cuenta que se trata de procesos bimoleculares limitados por la difusión (Fig 3.1C), estos resultados sugieren que existe algún tipo de pre-equilibrio conformacional en la población dimérica de E2c. Este equilibrio ha de ser previo a la colisión con ADN, es decir, E2c se encontraría en más de un estado conformacional. Una manera de probar esto sería encontrar condiciones en las cuales se pueda cambiar la proporción de las poblaciones conformacionales y evaluar luego la competencia entre éstas en la interacción cinética con ADN.

Si el experimento de pseudoprimer orden se realiza a menor temperatura (9 °C), donde estas velocidades de reacción no se ven afectadas (ver más abajo, Fig 3.9) los dos eventos de unión persisten en todas las concentraciones ensayadas, a diferencia de lo que ocurre a 25°C donde una de las colisiones deja de observarse. Si estas colisiones están asociadas a poblaciones de conformeros de E2c, la competencia entre ambos es menos abrupta a 9°C que a 25°C (Fig 3.2B), sugiriendo una población diferencial de los estados conformacionales. Este cambio está asociado con un pequeño, pero medible, cambio de elipticidad en el espectro de dicroísmo circular de E2c (Fig 3.2C). Tomados en conjunto, estos datos refuerzan la hipótesis que E2c se encuentra en un pre-equilibrio conformacional, y que éste puede ser afectado por el cambio de temperatura.

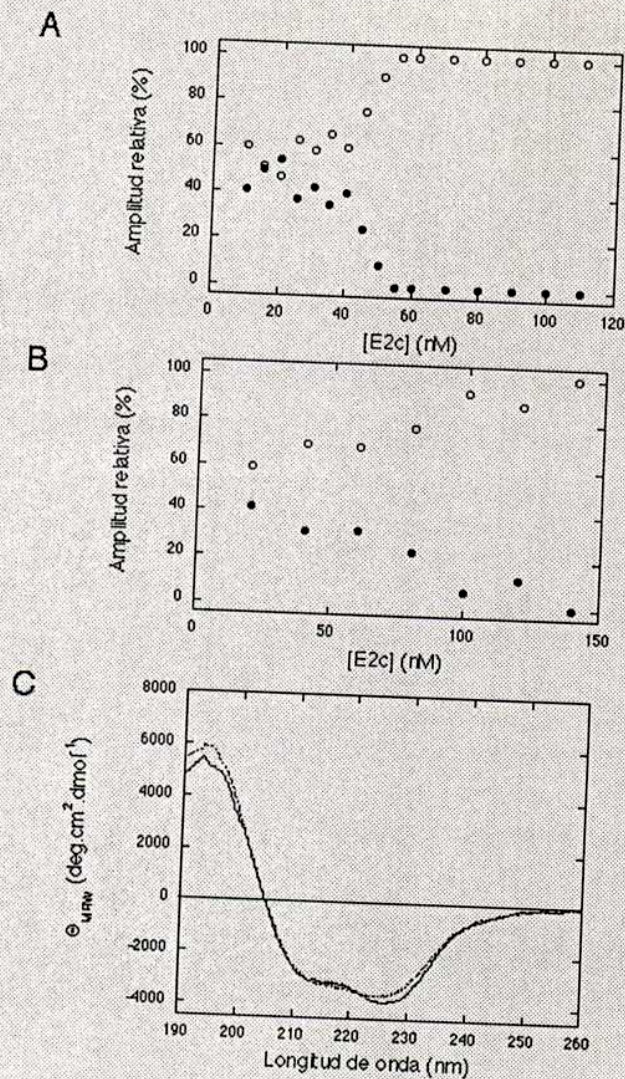


Figura 3.2: Población diferencial de confórmers de E2c. A) Amplitudes de las fases de asociación presentadas en la Fig 3.1, realizados a 25°C B) amplitudes de las mismas fases a 9°C. C) espectro de dicroísmo circular en el UV lejano de 10 μ M E2c en las condiciones de asociación, a 9°C (línea continua) y a 25°C (línea punteada).

En relación con lo presentado para el mecanismo de unión al equilibrio (capítulo 2), donde se observaron dos "modos" de unión diferentes (de relaciones molares 1:1 y 2:1 proteína:ADN), la existencia de conformómeros en la población de E2c puede estar asociada con la ocurrencia de estos "modos" diferenciales. Si se realiza un experimento de asociación de pseudo-primer orden respecto a E2c, utilizando ahora como ligando fijo un complejo 1:1 E2c:Fis35, se observa una sola fase de asociación de segundo orden (no mostrado). La k_{on} calculada para este evento es de $1.3 \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$, en concordancia con el evento cinético k_{on2} presentado en la tabla 3.1. Esto indicaría que sólo uno de los conformómeros responsable de la asociación a ADN libre es capaz de unir un complejo 1:1 proteína:ADN. Indicios sobre las bases estructurales de este conformómero, capaz de unir tanto ADN libre como un complejo 1:1 E2c-s35, serán discutidos en mayor profundidad en el capítulo 6.

COLISIÓN CON ADN NO-COGNATO Y NO-ESPECÍFICO

Como se detalló en el capítulo anterior, E2c une con alta afinidad sitios relacionados a s35 pero que no son "cognatos", como el derivado de BVP-1 (sBVP). En un experimento de pseudoprimer orden en exceso de E2c, utilizando FI-sBPV como ligando, se distingue una fase mayoritaria con una k_{on} similar a la obtenida para el s35, y otra fase de amplitud minoritaria pero claramente bimolecular. Los valores obtenidos para las constantes cinéticas de segundo orden se muestran en la tabla 3.1.

Utilizamos como ADN inespecífico los mismos oligonucleótidos descriptos para los experimentos al equilibrio. Para Iset35 se distinguen también dos eventos de colisión, siendo el más rápido de magnitud similar a la de los oligonucleótidos específicos. Esto es evidencia de que, aun para ADN inespecífico, las colisiones están limitadas por la difusión (tabla 3.1). Las disociaciones de estos complejos no pudieron ser cuantificadas con precisión, no obstante resulta evidente que k_{off} presenta un valor elevado, cercano a 100 s^{-1} . Para el sitio inespecífico sEBNA, vuelven a distinguirse dos fases de asociación, pero estas son 4 órdenes de magnitud más lentas que las detectadas para los sitios anteriores. Dado que E2c no reconoce ninguna secuencia en este oligonucleótido y que el sitio inespecífico Iset35 es unido a velocidades comparables a los específicos, la diferencia en la asociación puede deberse a restricciones de orientación conformacionales, algo que no ocurriría en un duplex con secuencia al azar como Iset35⁸⁵.

El hecho de que la unión de todos los oligonucleótidos presente dos eventos cinéticos de segundo orden, que además compiten entre sí, es evidencia independiente de que es E2c la que se encuentra en más de una conformación, y no el ADN FI-s35. Por tanto, es plausible descartar que la unión descrita con s35 se deba a un artefacto del oligonucleótido o al motivo fluorescente asociado.

Tabla 3.1: Parámetros cinéticos de los "complejos de encuentro" y de las disociaciones de "complejos finales" de E2c y diferentes oligonucleótidos.

Olgo	k_{on2} (M ⁻¹ s ⁻¹)	k_{on1} (M ⁻¹ s ⁻¹)	k_{-1} (s ⁻¹)	k_{-2} (s ⁻¹)	$k_{-1}(s^{-1})^a$	$k_{-2}(s^{-1})^a$
site35	1.4x10 ⁹	0.6x10 ⁹	6.3	ND	0.018	0.004
siteBOV	1.2x10 ⁹	6.0x10 ⁸	ND	0.07	0.037	0.009
siteEBNA	3.6x10 ⁸	1.8x10 ⁸	0.12	0.032	0.57	0.027
iset35	1.5x10 ⁹	1.1x10 ⁸	107	ND	1.73	0.34

a- medidos como la disociación del complejo final

CINÉTICA DE DISOCIACIÓN DE LOS COMPLEJOS E2c-ADN

Para la medición de las constantes cinéticas de disociación realizamos experimentos de competencia y captura entre Fl-s35 y s35. Fl-s35 fue incubado con un equivalente de E2c, lo cual produce un aumento en la intensidad de la fluorescencia del motivo FITC en paralelo a la formación del complejo proteína-ADN. Luego de incubar este complejo por 20 min para asegurar que se alcanzó el estado de equilibrio, se agregó un exceso de 50 veces s35 sin marca fluorescente, provocando la caída de la intensidad de fluorescencia a los niveles basales, lo cual indica la completa disociación del complejo Fl-s35:hE2c (Fig 3.3A). Para una determinación precisa de los parámetros que describen la disociación, el mismo experimento fue realizado en un aparato de flujo detenido. La caída de la intensidad de fluorescencia de Fl-s35 puede ser descrita por un decaimiento biexponencial, adjudicando el 70% de la amplitud a una fase de 0.004 s⁻¹ , y el 30% a una fase de 0.018 s⁻¹ (Fig 3.3A, tabla 3.1). Un experimento control muestra que si en el mismo ensayo se omiten la proteína o el ADN competidor, la intensidad de fluorescencia permanece constante (no mostrado).

Han sido descritos mecanismos por los cuales proteínas unidas a una molécula de ADN pueden "saltar" o "transferirse" a otra molécula de ADN sin que sea necesaria la disociación previa del primer complejo ^{106, 111}. El hecho de utilizar oligonucleótidos pequeños (18pb) minimiza tales efectos, mas no pueden ser descartados *a priori*. Si estos ocurrieran, la velocidad de transferencia debe depender de la concentración de ADN utilizado como desplazante. La figura 3.3B muestra que las velocidades de disociación del complejo E2c-Fls35 son independientes de la concentración de ADN utilizado como competidor, descartando que alguno de éstos mecanismos de "difusión facilitada" esten ocurriendo durante el evento de disociación medido, y confirmando que las velocidades obtenidas reflejan la disociación intrínseca del complejo proteína:ADN.

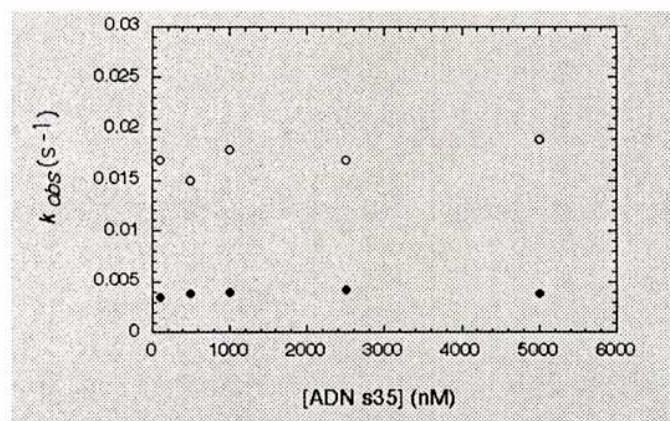
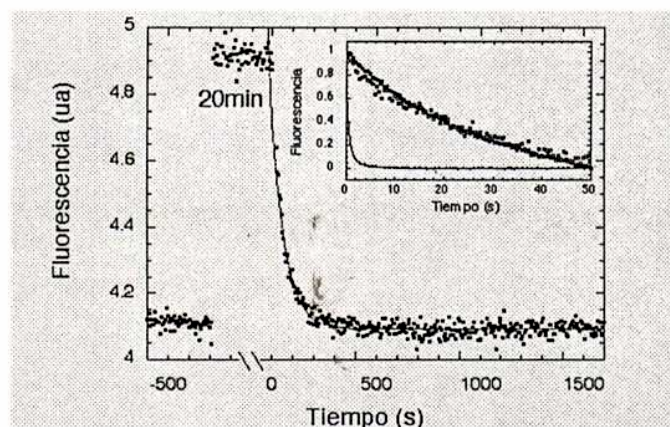


Figura 3.3: Cinética de disociación de los complejos E2c:ADN. A) 50nM FI-s35 fue mezclado con 50nM E2c y luego de 20 min se agregó un exceso de 50 veces s35, mientras se colectó la intensidad de fluorescencia de FITC. Inserto: curvas obtenidas por flujo-detenido para la disociación de los complejos E2c y diferentes oligonucleótidos: sBPV (círculos negros), sEBNA (círculos blancos), iset35 (línea). Todos las curvas ajustan a dos exponenciales, con los parámetros mostrados en la tabla 3.1. B) Los complejos 1:1 E2c:FI-s35 50nM fueron mezclados con diferentes concentraciones de s35. k_{obs} 1 (círculos blancos), k_{obs} 2 (círculos negros). Experimentos preeliminarios realizados al equilibrio mostraron que el agregado de un exceso molar de 25 veces s35 resulta suficiente para observar este desplazamiento.

Ensayamos la cinética de disociación para los complejos de E2c y diferentes oligonucleótidos. Para todos los casos, la disociación presenta un comportamiento bifásico sugiriendo que, como en el caso de la asociación, la disociación del complejo E2c:ADN ocurre por más de una ruta (inserto Fig 3.3A, tabla 3.1). Los valores de las constantes son mayores cuando se trata de oligonucleótidos no específicos, iset35 y EBNA, y solamente se duplican para el sitio no-cognato sBPV (tabla 3.1). Teniendo en cuenta la similitud de los valores medidos para la asociación, estas mediciones apuntan a los eventos de disociación como responsables de la capacidad de discriminación de secuencia que presenta E2c. En estos casos, es de tener en cuenta que no es posible comparar directamente las velocidades y amplitudes relativas para las fases de cada oligonucleótido, dado que las señales fluorescentes pueden no estar relacionadas a la concentración relativa.

La unión entre E2c y s35 presenta un cambio en la intensidad de fluorescencia de los residuos Trp localizados en el centro del barril- β (Fig 2.2, capítulo 2). Buscando corresponder los rearrreglos estructurales y las señales espectroscópicas que podemos detectar, realizamos

CINÉTICA DE ASOCIACIÓN Y REARREGLOS CONFORMACIONALES EN E2C

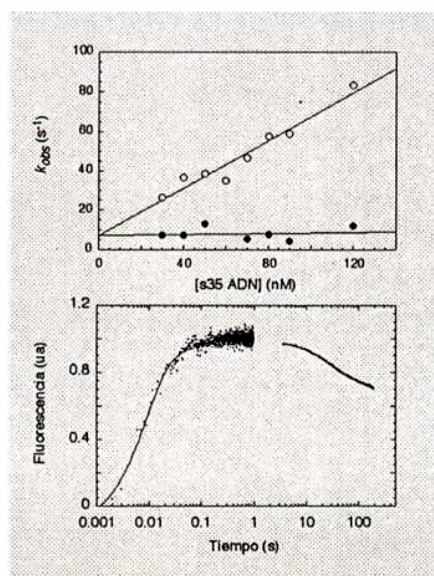


Figura 3.4: Cinética de la asociación E2c:s35 seguida por fluorescencia de Trp. A) pseudoprimer orden a 25 nM E2c. Los datos ajustaron a dos exponenciales: k_{on1} (círculos blancos), k_2 (círculos negros). B) 300 nM E2c y 300 nM s35 fueron mezclados en un equipo de flujo detenido y el cambio de fluorescencia de Trp colectado en dos escalas temporales. Las líneas muestran ajuste biexponencial de 0 a 1 s (k_{on1} , k_2), y monoexponencial de 1 a 200 s (k_3).

experimentos cinéticos de unión entre E2c y s35, registrando el cambio de intensidad de la fluorescencia intrínseca de Trp. Las curvas preliminares realizadas con la proteína salvaje presentaron decaimientos multiexponenciales que dificultan el análisis, por lo que optamos por la medición de las señales asociadas a la unión de ADN de la mutante W339F, descrita anteriormente (ver capítulos 2 y 7). En condiciones de pseudoprimer orden respecto a E2c, en exceso de ADN y a una concentración fija de 25 nM proteína, los datos ajustan a dos procesos exponenciales simples. De ellos, sólo uno depende de la concentración de ADN (Fig 3.4A). La k_{on} obtenida es de $0.6 \times 10^9 M^{-1} s^{-1}$, en perfecta concordancia con la fase más lenta (k_{on1}) de las descritas para el caso de la asociación con FI-s35 (Fig 3.1B). La fase de $8 s^{-1}$, es independiente de la concentración de s35, por lo que es asignable a algún tipo de rearreglo conformacional del dímero de E2c. Este rearreglo involucra un cambio en el entorno electrónico de los Trp, localizados en el centro del barril β -dímero. La k_{off} del evento de unión es cercana a $10 s^{-1}$, permitiéndonos estimar una K_D de 20 nM para este "complejo de encuentro". A diferencia de lo observado en exceso de proteína sobre el FI-s35, no se detectaron otros eventos de colisión bimolecular. Esto puede deberse a que el conformero responsable no presente un cambio de fluorescencia intrínseca asociado a la unión, o bien que el evento no sea detectado dada la muy menor sensibilidad de la sonda fluorescente utilizada (Trp) frente a FITC.

La estructura cristalográfica del complejo E2c-ADN muestra que existe una extensa red de contactos cercanos entre las cadenas laterales y los motivos de las bases nitrogenadas^{39 34} (Fig 4.1, capítulo 4), y las comparaciones con las estructuras de E2c libre sugieren que deben existir cambios conformacionales asociados a la unión³⁹. Es poco probable que todos estos eventos sucedan en la escala temporal de 100 ms, el tiempo medio de la fase de $8 s^{-1}$. Buscando señales fluorescentes que evidencien fases más lentas, aumentamos la concentración de E2c con el fin de acelerar y separar temporalmente el evento de colisión inicial, y a la vez aumentar el nivel de detección de las señales fluorescentes de los Trp. A 300 nM E2c se distinguen las fases de unión y el rearreglo de $8 s^{-1}$ y, en otro rango temporal, un evento posterior (Fig 3.4B). Esta fase, cuya k_{obs} es de $0.04 s^{-1}$ no depende de la concentración de s35 y, a diferencia de las fases más rápidas, presenta una disminución de la intensidad de fluorescencia (Fig 3.4B). El tiempo medio asociado a este evento es de 17 sec.

Se registraron también los cambios fluorescentes de los Trp cuando E2c une el sitio específico (pero no-cognato) sBPV. Al igual que para s35, se registran dos fases independientes de la concentración de ADN. Estas están en el mismo orden temporal que las descritas para la unión de s35, y las intensidades relativas de los cambios fluorescentes están en perfecta concordancia. Esto nos permite inferir que las diferencias encontradas para la unión al equilibrio de E2c y éstos oligonucleótidos, no se deben a diferencias asignables a estas etapas cinéticas sino a otros eventos, presumiblemente a la cinética de disociación (Fig 3.3, tabla 3.1).

EFFECTO DE LA FUERZA IÓNICA EN LAS FASES CINÉTICAS

El evento de asociación detectado por fluorescencia de Trp es altamente dependiente de la fuerza iónica (Fig 3.5A). Esto puede corresponder a dos efectos diferentes en tanto que las cargas modifican la constante dieléctrica del entorno y además pueden neutralizar las cargas expuestas de las macromoléculas. Para el caso de asociaciones proteína-ADN, los contraiones Na^+ que neutralizan los grupos fosfato son desplazados ante la unión de la proteína, actuando entonces como reactantes ya que:

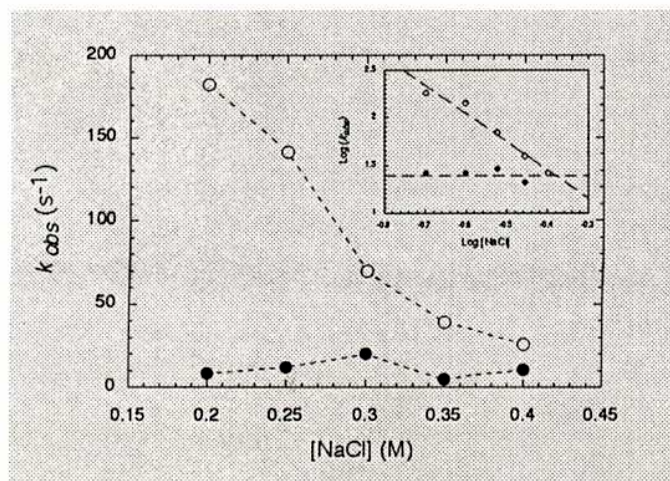


siendo $[\text{P}]$, $[\text{ADN}]$ y $[\text{Na}^+]$ las concentraciones de proteína, ADN e ión sodio, respectivamente. Utilizando la relación empírica: ^{17, 75}

$$d \log k / d \log [\text{Na}^+] = -0,88 m$$

es posible calcular de la pendiente del gráfico doble logarítmico que alrededor de 6 contraiones Na^+ son desplazados de la interfase de unión del "complejo de encuentro" E2c-ADN. Este valor debe ser tomado como aproximado, estimando un error del 50%, ya que el cálculo no incluye el posible efecto de la fuerza iónica sobre la constante dieléctrica del medio, lo que podría "apantallar" las cargas en una asociación limitada por la difusión y asistida electrostáticamente, como es el caso de E2c-ADN. De todas formas, el valor está en el orden del número estimado por los puentes salinos formados al equilibrio entre E2c y ADN, según los estudios cristalográficos ^{34, 35}. Esto muestra además que el desplazamiento de contraiones ocurre durante el primer evento de unión. La segunda fase (k_2) no depende de la fuerza iónica del entorno, sugiriendo que no hay cargas expuestas asociadas a la ocurrencia de este rearrreglo conformacional en E2c.

Figura 3.5: Efecto de la fuerza iónica en la cinética de interacción E2c:s35
25nM E2c fue mezclado con 70nM s35 a diferentes concentraciones de NaCl. La fluorescencia de Trp fue colectada y los datos ajustados a un decaimiento biexponencial. Inserto: Gráfico doble logarítmico y ajuste lineal de la dependencia de la fuerza iónica de las mismas fases.



EXCLUSIÓN DE SOLVENTE DE LA INTERFASE

La formación de un complejo consolidado proteína-ADN específico requiere la formación de una extensa red de interacciones cercanas, conjuntamente con la remoción de solvente de la interfase de unión. No habiendo una manera directa de observar tales eventos, hicimos uso de la sonda fluorescente ANS. Esta sonda es particularmente útil para medir cambios conformacionales en proteínas, dado el elevado cambio en el rendimiento cuántico asociado al cambio en el microentorno del fluoróforo ³². Se ha descrito que, en términos generales, el ANS es capaz de unirse a proteínas que unen ADN, con la particularidad de hacerlo en las zonas de unión de ADN, ya que en su mayoría tienen densidad de carga positiva y un carácter parcialmente hidrofóbico ⁶¹.

Datos previos mostraron que E2c une ANS o bis-ANS en al menos dos sitios, uno de los cuales es desplazado por la unión de ADN ^{35, 73}. Dado el pequeño tamaño relativo del fluoróforo y sus características químicas, suponemos que el ANS podría quedar "atrapado" en la interfase de unión, conjuntamente con el solvente. La remoción del ANS, entonces, podría reflejar un evento de expulsión de solvente de la interfase de interacción. E2c fue incubado con concentraciones saturantes de ANS, dando un incremento en la intensidad de fluorescencia de ANS, como era esperado. Al mezclar la muestra con un equivalente de ADN s35 (también en ANS), se observó una disminución de la intensidad de la fluorescencia, indicando el desplazamiento del fluoróforo del sitio de unión (Fig 3.6A). Este decaimiento ajusta a un decaimiento monoexponencial simple con $k_{ANS} = 0.04 \text{ s}^{-1}$, valor que es independiente de la concentración de reactantes utilizados (Fig 3.6A inserto). Experimentos de pseudoprimer orden respecto a E2c llevados a cabo en presencia de las mismas concentraciones de ANS, muestran que el fluoróforo no afecta las constantes de asociación de E2c:FI-s35 (Fig 3.6B). Dado que la velocidad de disociación espontánea del ANS de E2c es un orden más lenta (0.003 s^{-1} , no mostrado), interpretamos la fase detectada como la expulsión del ANS de la interfase proteína-ADN. Uno de los rearrreglos evidenciados por fluorescencia de Trp (k_3 , Fig 3.4) tiene una velocidad idéntica, sugiriendo que la expulsión de solvente ocurre concomitantemente con un cambio conformacional en E2c, algo no del todo sorprendente si consideramos las consecuencias estructurales esperables para la desolvatación parcial de una interfase extensa.

Si el mismo experimento se realiza desplazando el ANS con ADN no-cognato (sBPV), se observa una fase con $k_{ANS} = 0.032 \text{ s}^{-1}$, muy cercana a la obtenida para s35. En cambio, los oligonucleótidos inespecíficos sEBNA e Iset35 no fueron capaces de desplazar ANS, incluso cuando fueron ensayados a concentraciones 10 veces superiores a las K_D 's respectivas. Estos resultados muestran que, a pesar de unirse rápidamente a E2c (tabla 3.1), los complejos E2c-ADN inespecífico no muestran rearrreglos que incluyan la expulsión de solvente de la interfase, en plena concordancia con otros modelos de reconocimiento proteína-ADN. Estos proponen que la unión con ADN inespecífico es mayoritariamente electrostática, no incluye contactos cercanos, y las interfases de unión se mantendrían hidratadas ^{24, 44, 74}. En este sentido, lo presentado representa una de las primeras mediciones experimentales directas de tales eventos.

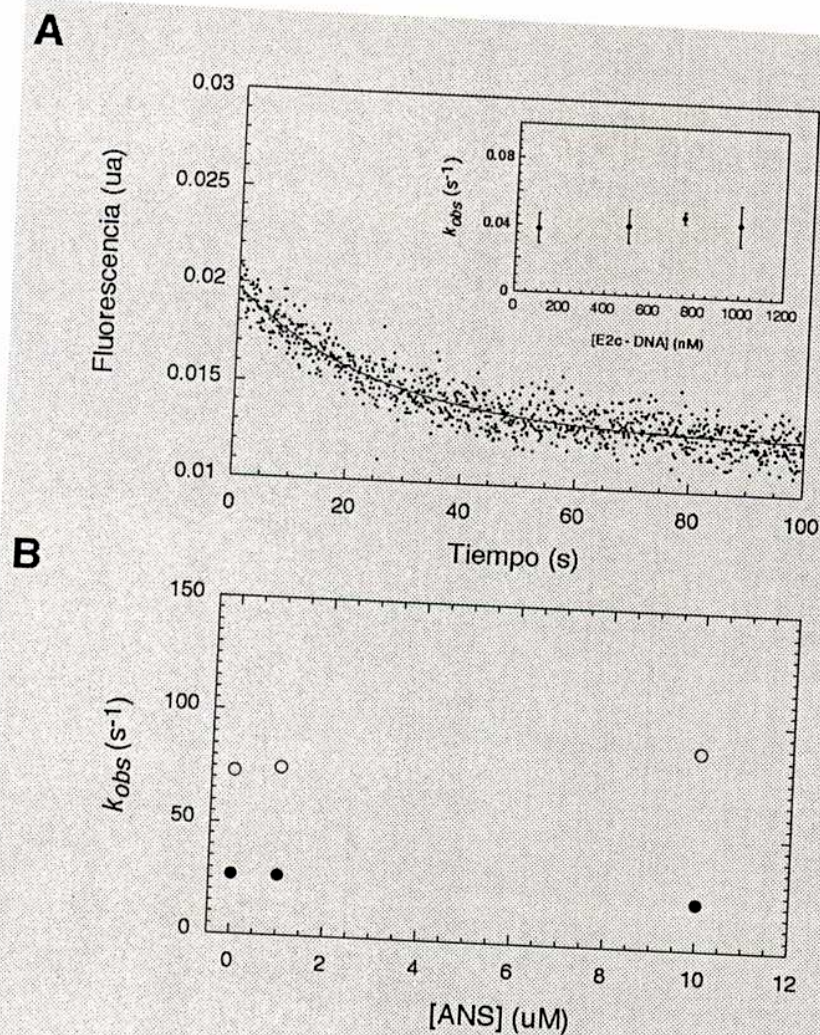


Fig 3.6: Desplazamiento de ANS de la interface de unión. A) $1\mu M$ E2c fue pre- incubado con $50\mu M$ ANS y se mezcló con $1\mu M$ s35 en $50\mu M$ ANS. Se muestran los datos de fluorescencia de ANS de un experimento típico. Inserto: dependencia de la velocidad con la concentración de reactantes. B) $5nM$ Fl-s35 fue mezclado con $40nM$ E2c en diferentes concentraciones de ANS. Se colectó la fluorescencia de FITC, y se grafican las constantes obtenidas del ajuste biexponencial.

IDENTIFICACIÓN DE RUTAS DE ASOCIACIÓN-DISOCIACIÓN PARALELAS

Los eventos de disociación de los complejos E2c:ADN presentan más de una fase fluorescente (tabla 3.1). Para determinar si estos eventos representan caminos de disociación paralelos o secuenciales, realizamos experimentos de doble-mezcla (double-jump). En un aparato de flujo detenido mezclamos FI-s35 y E2c en cantidades equimolares. Según lo caracterizado en las secciones anteriores podemos calcular que a una concentración de 100 nM el $t_{1/2}$ de los eventos de colisión es de 12 mseg. A tiempos de incubación variables, de 0.01 a 100 seg, competimos la mezcla con un exceso de s35 mientras registramos el decaimiento de la fluorescencia del motivo FITC del oligonucleótido (Fig 3.3). Este arreglo de doble-mezcla permite medir los eventos de disociación de las poblaciones de intermediarios que se forman a los diferentes tiempos de incubación, variables. Encontramos que el 70% de la amplitud del cambio fluorescente corresponde a 2 fases lentas, de 0.0037 s^{-1} y 0.0011 s^{-1} , coincidentes con las fases encontradas al disociar el "complejo consolidado" E2c:FI-s35 (tabla 3.1 y Fig 3.7A). Sorprendentemente, vimos que estas fases están presentes incluso desde los 10 mseg de iniciada la reacción de unión (Fig 3.7A), y que éstas se observan durante todo el tiempo de incubación. Esto es indicativo de que pueden formarse cantidades detectables del "complejo final" a tiempos cortos de iniciada la reacción de asociación, probablemente en la velocidad límite impuesta por la difusión.

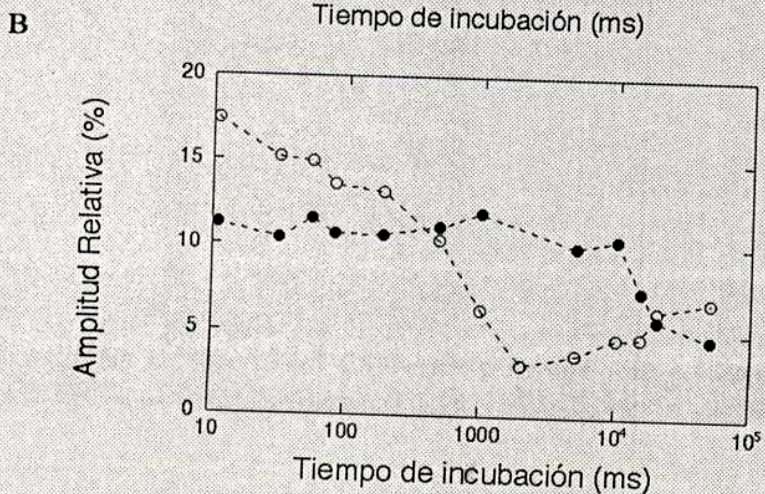
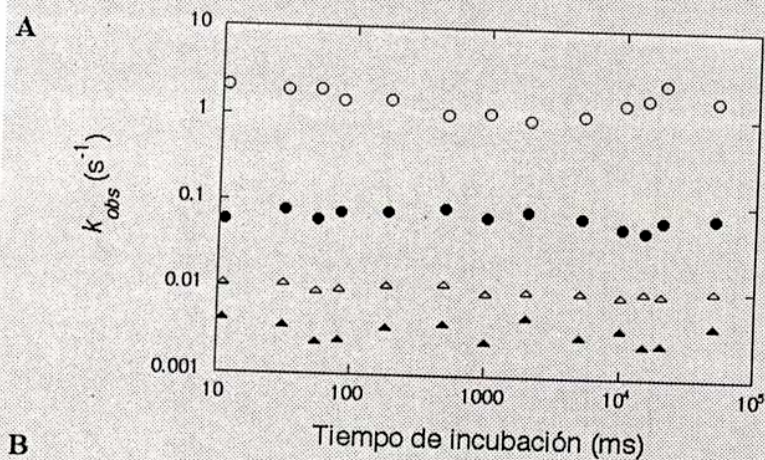
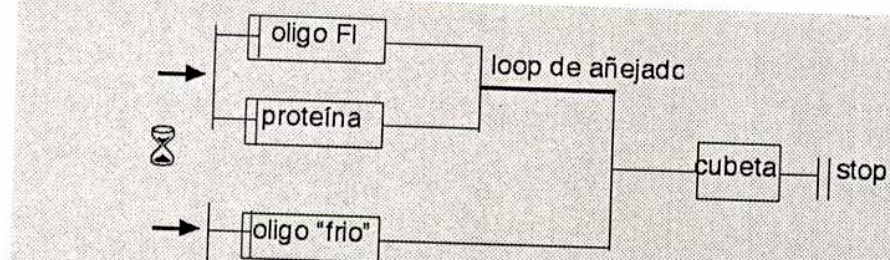
Además pudimos identificar otras dos fases de disociación, con constantes de 1 s^{-1} y 0.07 s^{-1} (Fig 3.7A) cuyas amplitudes relativas disminuyen a medida que transcurre el tiempo de incubación, y desaparecen a los 2 y 20 sec respectivamente (Fig 3.7B). Esto sugiere que estas dos fases son eventos secuenciales, siendo la primera de ellas coincidente con la disociación del complejo de encuentro, reportado en el experimento de pseudoprimer orden (k_1 , Fig 3.1B), y la segunda es atribuible a la reacción reversa del primer rearreglo conformacional (k_2 , Fig 3.4A).

Estos resultados son evidencia firme de la existencia de dos caminos paralelos de asociación entre E2c y FI-s35, en concordancia con la detección de dos rutas paralelas de unión descritas por los experimentos de pseudoprimer orden (Fig 3.1, Fig 3.2). La estrecha ventana temporal del análisis, y la invariancia de las amplitudes relativas de k_{off1} y k_{off2} , sugiere además que las poblaciones finales a las cuales llega el complejo E2c-s35 están en rápido equilibrio. Por lo tanto, ambas poblaciones deben estar estrechamente relacionadas estructuralmente, y la barrera energética que las separa debe ser pequeña.

Figura 3.7: Experimentos de asociación-disolución por doble-mezcla. 100 nM E2c fue mezclado con 100nM FI-s35 y luego de un "tiempo de incubación" variable, la mezcla fue competida con 2.5µM s35. Los perfiles de fluorescencia de FITC luego de la segunda mezcla fueron ajustados a decaimientos exponenciales y se muestran: En la parte superior, un esquema del arreglo experimental para la doble-mezcla.

A) las constantes obtenidas: k_{obs} 1 (círculos blancos), 2 (círculos negros), 3 (triángulos blancos), 4 (triángulos negros).

B) se muestran las amplitudes correspondientes a las fases 1 y 2 indicadas en A. Las amplitudes de las fases 3 y 4 permanecen constantes.



Experimentos de doble-mezcla como el presentado en la figura 3.7 fueron llevados a cabo con los otros oligonucleótidos. Para el sBPV se identificaron 3 fases: i) un evento rápido con $k_{obs1} = 2 \text{ s}^{-1}$ cuya amplitud desaparece a los 2 sec, y la asignamos a la disociación del complejo de encuentro, ii) una con $k_{obs2} = 0.008 \text{ s}^{-1}$, coincidente con la fase mayoritaria encontrada para la disociación del complejo consolidado (tabla 3.1) que, como para el s35, esta presente desde los primeros tiempos de incubación; iii) una con $k_{obs3} = 0.004 \text{ s}^{-1}$, coincidente con la otra de las encontradas para la disociación del complejo consolidado, que tiende a desaparecer a los 25 sec.

Para el caso del ADN inespecífico FI-iset35 se detectaron dos fases coincidentes con la disociación del complejo final (tabla 3.1), estando presentes desde los tiempos más cortos de incubación, con constantes de velocidad de $2,0$ y $0,5 \text{ s}^{-1}$. El hecho de que no se observen más eventos fluorescentes para esta asociación concuerda con lo hipotetizado anteriormente acerca del modo de unión a ADN inespecífico.

Los resultados presentados muestran que la asociación entre hE2c y ADN (específico o no) puede darse por al menos dos rutas paralelas. Para el caso de ADN específico, una de las rutas incluye el pasaje a través de dos intermediarios y la otra, extremadamente rápida, puede ser representada como de "dos estados". En la figura 3.8 se presenta un esquema del modelo que proponemos para las reacciones cinéticas de unión E2c-ADN específico.

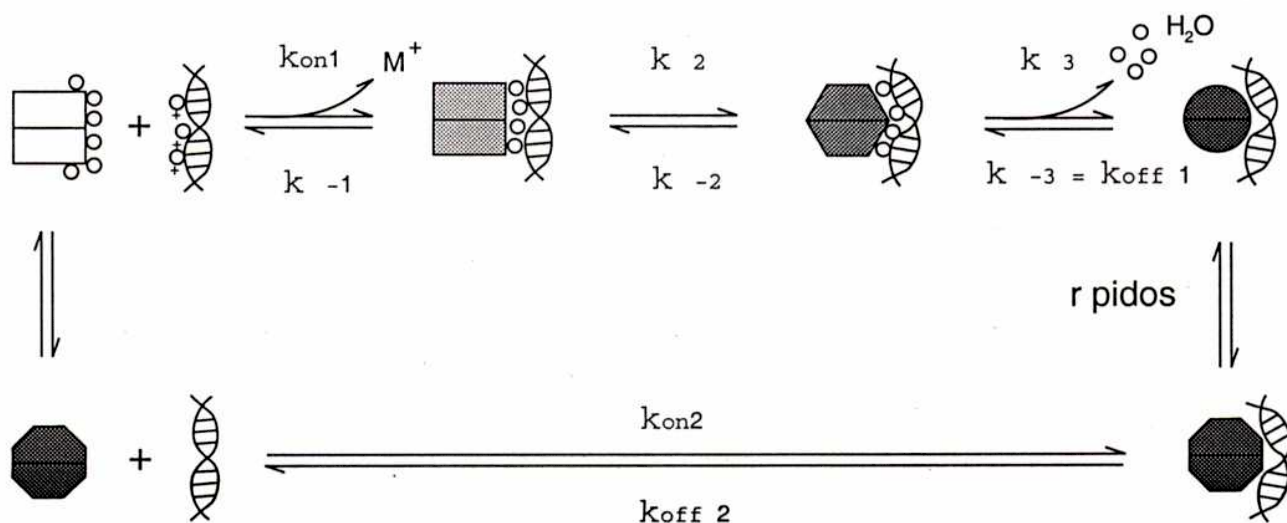


Figura 3.8: Esquema de la interacción específica E2c:ADN. E2c existe en dos poblaciones conformacionales, cada una de las cuales une ADN por rutas paralelas. En la parte inferior: "ruta de dos estados"; en la parte superior "ruta de intermediarios". Los valores de las constantes cinéticas se muestran en la tabla 3.1.

EFFECTO DE LA TEMPERATURA SOBRE LOS EVENTOS DE ASOCIACIÓN: CARAC- TERIZACIÓN TERMODINÁMICA DEL TSE DEL EVENTO DE COLISIÓN INICIAL

La dependencia de los eventos cinéticos con la temperatura es una herramienta fundamental para obtener información acerca de los intermediarios y de las barreras de los estados de transición que los separan ²⁵. Para evaluar el efecto de la temperatura sobre la reacción de unión, elegimos seguir los cambios de las señales de Trp de E2c (W339F), ya que se evidencia una sola ruta de unión y, simultáneamente, provee información acerca de los rearrreglos conformacionales de los intermediarios.

Realizamos experimentos de unión en pseudoprimer orden respecto a s35, a diferentes temperaturas, entre 9 y 35 °C (Fig 3.9A). Como es evidente en la figura y se detalla en el inserto, el efecto de la temperatura tanto en k_{on} o k_{-1} es muy pequeño, dentro de los errores asociados a la medición.

El hecho que la temperatura afecte en poca medida las reacciones de asociación impidió la extracción de los parámetros termodinámicos precisos, por lo que se realizó un experimento complementario. Medimos el valor de la constante cinética observada a una concentración fija de ligandos, en función de la temperatura (Fig 3.9B). El ajuste de los datos a la ecuación 7.2 (capítulo 7) muestra que la reacción se puede describir con una barrera de activación pequeña, esperable para un evento guiado entrópicamente (tabla 3.2). Esto esta en concordancia con los ensayos de fuerza iónica presentados en la Fig 3.5, donde se mostró que el evento de colisión involucra la liberación de contralones de la interfase, para lo cual que se espera una contribución mayoritariamente entrópica ^{47, 78}. Por lo tanto, los efectos de la temperatura parecen estar restringidos a aquellos relacionados con el movimiento y difusión de las macromoléculas en solución.

Otra de las contribuciones energéticas a la termodinámica de interacción es el cambio de capacidad calorífica a presión constante, ΔC_p . Las contribuciones del ΔC_p a la energía libre de los TSE estan dadas mayoritariamente por el *derretido* de los "clusters" de solvente que hidratan las superficies hidrofóbicas ²⁵. Si uno considera la diferencia de exposición al solvente

como la única fuente de ΔC_p ^{58, 75}, la linealidad del gráfico de Eyring ($\Delta C_p^\ddagger = 0$) sugiere que hay cambios mínimos en la exposición de superficie, tanto para la reacción partiendo desde las macromoléculas libres como desde el complejo de encuentro, al conjunto de estados de transición que los separa. Esto implicaría que la interfase del "complejo de encuentro" permanecería con un grado de solvatación cercano al de macromoléculas libres.

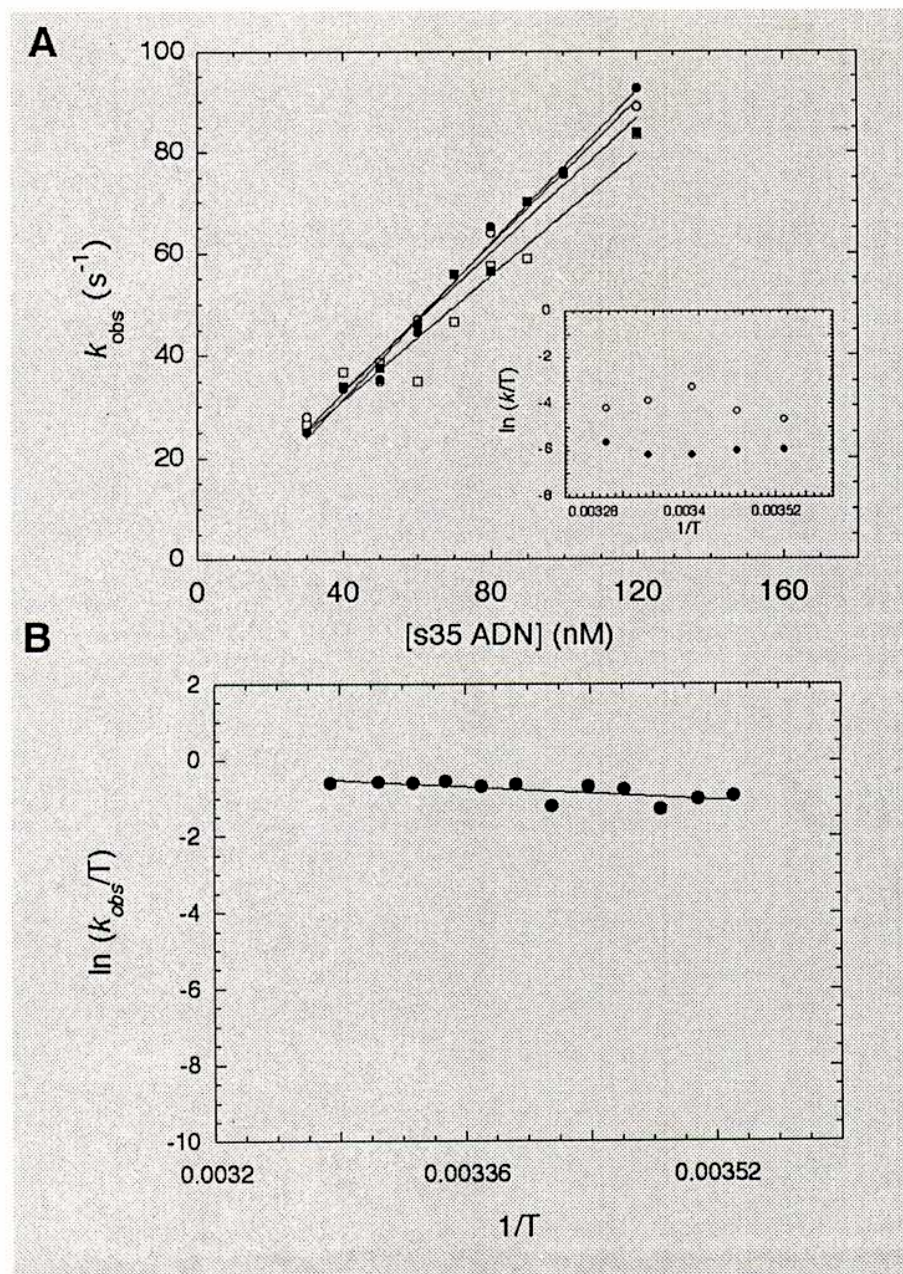


Figura 3.9: Efecto de la temperatura en la colisión inicial. A) 25nM E2c fue mezclado con diferentes concentraciones de ADN a diferentes temperaturas: 10°C (círculos negros), 15°C (círculos blancos), 20°C (cuadrados negros), 25°C (cuadrados blancos). La unión se identificó por la fluorescencia de Trp, como en la Fig 3.4. Inserto: gráfico de Eyring de la k_{on1} (círculos negros) y la k_{-1} (círculos blancos) calculadas. B) Gráfico de Eyring de la constante de asociación observada para el complejo de encuentro a 30nM E2c y 60nM s35.

Tabla 3.2: Parámetros termodinámicos calculados para las barreras de activación de los rearreglos conformacionales de la interacción E2c-ADN específico

	k_{on1}	k_2	K_3	k_{off2}	k_{off1}
$\Delta G_0^\ddagger$ (kcal mol ⁻¹)	4.8 ± 1.1	16.1 ± 3.6	19.2 ± 3.2	19.6 ± 0.1	20.4 ± 0.5
$\Delta H_0^\ddagger$ (kcal mol ⁻¹)	1.5 ± 0.3	21.5 ± 2.6	18.0 ± 3.0	11.0 ± 1.3	15.4 ± 7
$\Delta S_0^\ddagger$ (cal K mol ⁻¹)	-11.1 ± 1	18.1 ± 2.5	-4.0 ± 1.2	-28.9 ± 1.3	-17.0 ± 7
$\Delta Cp^\ddagger$ (kcal mol ⁻¹ K ⁻¹)	ND	ND	ND	-0.26 ± 0.1	-0.98 ± 0.4

ANÁLISIS DE LOS ESTADOS DE TRANSICIÓN ASOCIADOS A LOS REARREGLOS CONFORMACIONALES

La velocidad de los eventos de rearreglo detectados por los cambios en la intensidad de fluorescencia de los Trp de E2c (W339F) fue medida en el rango de temperaturas experimentalmente accesibles, entre 9 y 35 °C (Fig 3.10). Los gráficos de Eyring muestran que las velocidades aumentan a medida que aumenta la temperatura, indicando energías libres de activación positivas. Los ajustes a la ecuación 7.3 (capítulo 7) permiten estimar los valores de los parámetros termodinámicos de los conjuntos de estados de transición, presentados en la tabla 3.2. Los valores de la energía libre de los TSE ($\Delta G^\ddagger$) son altos comparados con el valor de referencia (la energía térmica a 25°C, RT = 0,6 kcal/mol) y con la diferencia energética total medida al equilibrio (-13.5 kcal/mol, Fig 2.5), definiendo "valles" de mínimos energéticos profundos donde caen los intermediarios.

La barrera de activación del primer rearreglo (k_2), tiene contribuciones entálpicas desfavorables y entrópicas favorables, siendo esperados valores positivos de $\Delta S^\ddagger$ cuando aumentan los grados de libertad rotacionales y traslacionales internos de los reactantes (tabla 3.2). En esta instancia no es posible disectar las contribuciones de cada macromolécula, proteína o ADN, a las energías de activación presentadas. Ha sido descrito (en los capítulos precedentes y en ^{77, 78, 79}) que tanto E2c como el oligonucleótido presentan algún grado de plasticidad estructural, y que ambos reactantes inducen mutuos cambios conformacionales al interactuar. Por tanto, el complejo proteína:ADN debe ser interpretado como una entidad termodinámica nueva y diferente, más allá de las contribuciones independientes de las partes.

La mayor contribución a la energía libre de activación del segundo rearreglo (k_3), es entálpica, cuyo origen puede asignarse a la formación de interacciones de van der Waals y uniones de hidrógeno, presumiblemente en la interfase de asociación (tabla 3.3). Para que estas uniones cercanas tengan lugar, las esferas de hidratación de las macromoléculas deben reorganizarse y, como se mostro en la Fig 3.6, esto ocurre en el mismo rango temporal que k_3 . Tomados en conjunto, estos datos sugieren que la "lectura directa" de las bases del ADN se consolidaría durante este evento.

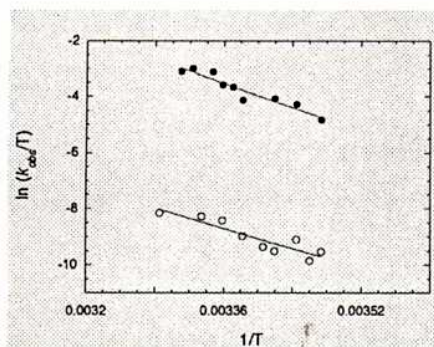


Figura 3.10: Efecto de la temperatura a los rearrreglos conformacionales de E2c-ADN. Gráfico de Eyring de las constantes observadas para las fases fluorescentes Trp, independientes de la concentración, a 100nM E2c y 300nM s35. k_2 (círculos negros), k_3 (círculos blancos).

EFFECTO DE LA TEMPERATURA SOBRE LA DISOCIACIÓN DEL COMPLEJO E2C:S35

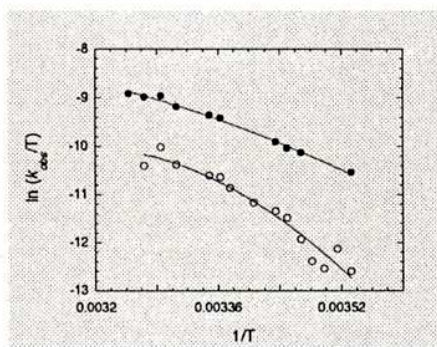


Figura 3.11: Efecto de la temperatura sobre la disociación de E2c-ADN. Gráfico de Eyring de las constantes observadas para las disociaciones del complejo E2c:Fls35, medidas como en la Fig 3.3. A todas las temperaturas, los decaimientos ajustaron a dos exponenciales: k_{off1} (círculos negros), k_{off2} (círculos blancos).

Para ninguno de los rearrreglos descritos se detectó un componente $\Delta C_p^\ddagger$ importante, algo inesperado dado que un fuerte ΔC_p negativo se midió por estudios de calorimetría de titulación isotérmica de la unión E2c:s35 al equilibrio²⁰. Esto podría indicar que la contribución al $\Delta C_p^\ddagger$ esta presente en rearrreglos conformacionales subsiguientes, no detectados por cambios en la intensidad de fluorescencia de Trp. Otra posibilidad es que, dado que estamos midiendo la transición desde los reactivos libres pasando por los intermediarios y los estados de transición, es posible que el componente $\Delta C_p^\ddagger$ se evidencie en la reacción reversa, es decir, en la disociación del complejo final.

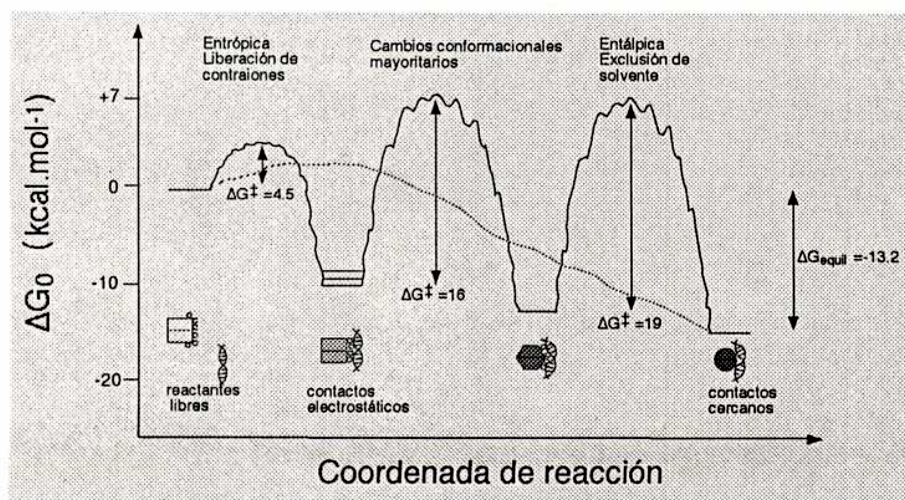
Con el objeto de medir el efecto de la temperatura sobre los eventos de disociación utilizamos el ensayo descrito para la disociación del complejo E2c:Fl-s35 (Fig 3.3), donde se midió la disociación por el agregado de un exceso de ADN sin fluoresceinar. A todas las temperaturas ensayadas se detectaron dos fases fluorescentes, correspondientes a las rutas de disociación paralelas descritas (Fig 3.11). Ambas reacciones tienen altas energías libres de activación, en el mismo rango que las reacciones en el sentido inverso (k_3), evidenciando que son los pasos limitantes de las rutas de disociación (tabla 3.3). Las energías de activación tienen contribuciones entrópicas y entálpicas desfavorables, lo cual es esperado para la redistribución de la red de uniones en la interface y las entropías rotacionales y vibracionales asociadas al proceso. Se detectó además un fuerte componente, de signo negativo, en el cambio de capacidad calorífica ($\Delta C_p^\ddagger$, tabla 3.3). Esto puede interpretarse como un "ocultamiento" de superficie accesible al solvente ocurriendo en esta etapa. Cabe remarcar que esto está en concordancia con el evento de exclusión de solvente de la interfase, medido por la expulsión de ANS, postulado justamente para esta etapa cinética (Fig 3.6).

EFFECTO DE LA TEMPERATURA SOBRE LA RUTA DE ASOCIACIÓN DE DOS ESTADOS

Como fue descrito, la unión de E2c a ADN puede darse por una ruta caracterizada como de "dos estados" (Fig 3.1, Fig 3.7). A fin de estudiar el efecto de la temperatura sobre esta vía, intentamos realizar curvas de asociación a diferentes temperaturas utilizando el oligonucleótido FI-s35. Los efectos observados no pudieron ser interpretados dada la superposición entre la variación de las constantes y las amplitudes con las variaciones del rendimiento cuántico de la fluorescencia debidas solamente al cambio de temperatura. Sumado a esto, tendríamos que considerar también el efecto de la temperatura sobre la fluctuación de las poblaciones conformacionales de E2c (Fig 3.2), por lo que la cantidad de variables asociadas imposibilitan un análisis detallado.

No obstante, podemos efectuar algunas observaciones indirectas. Como se describió, la velocidad de asociación de esta ruta está limitada por la difusión (Fig 3.1) y, considerando que ésta es análoga a la presentada para el camino de unión con intermediarios (Fig 3.9), la barrera de activación para esta unión debe ser también pequeña. En este sentido, los experimentos de doble-mezcla muestran que a 25°C se puede alcanzar una población de "complejo final" en tiempos cortos, en concordancia con una barrera de activación de baja energía libre, en la dirección de la formación del complejo. Por otro lado, la barrera de disociación atribuida a esta ruta es alta (Fig 3.11, tabla 3.2), y teniendo en cuenta que la reacción puede ser planteada como de "dos estados", con una alta diferencia energética total, se corresponde con la asignación de una pequeña barrera de activación al proceso de asociación (Fig 3.12).

Figura 3.12: Representación gráfica de los cambios termodinámicos asociados a la interacción E2c-ADN. Los valores corresponden a las energías libres de activación estándar a 25°C y 1M, mostrados en la tabla 3.2. Para el "complejo de encuentro", las 3 líneas representan los valores de la unión de E2c y diferentes oligonucleótidos. La línea punteada muestra el perfil de reacción para la ruta de "dos estados", y la línea llena la de la ruta de "dos intermediarios".



RELACIÓN DE LAS MEDICIONES CINÉTICAS CON LA MEDICIONES AL EQUILIBRIO

Teniendo en cuenta las mediciones energéticas al equilibrio, conjuntamente con los resultados presentados en este capítulo, podemos construir esquemas de reacción para los caminos cinéticos de unión E2c-ADN (Fig 3.12). Para la ruta descrita como de dos estados, la relación de las velocidades de asociación y disociación nos permiten estimar una constante de disociación al equilibrio (K_D) de 1.5×10^{-10} M, en estrecha concordancia con la medición al equilibrio, 2×10^{-10} M (capítulo 2). Para el camino con dos intermediarios, la constante de disociación calculada para los equilibrios concatenados es de 0.04 nM, que puede considerarse muy aceptable teniendo en cuenta la complejidad del camino cinético y los errores asociados a las mediciones.

Según el modelo del mecanismo cinético de unión E2c:s25 presentado (Fig 3.8), existen dos caminos de unión paralelos. Estas rutas parten de un equilibrio conformacional de E2c, preexistente a la unión de ADN. Los "confórmers" planteados no deben ser interpretados como dos estructuras proteicas rígidas y diferentes, sino como poblaciones conformacionales dinámicas, como *ensembles* conformacionales separados por una barrera energética. De la dependencia de las velocidades de colisión con la concentración de E2c, se estimaron las constantes cinéticas de segundo orden (k_{on}), utilizando para el cálculo los valores de concentración total de E2c. Planteada la existencia de *ensembles* conformacionales, debemos reinterpretar los resultados, ya que la concentración de E2c total esta dada por la suma de la concentración de ambas poblaciones de confórmers:

$$[E2c] = [confA] + [confB]$$

A su vez, la constante de velocidad observada para cada colisión varía según:

$$k_{obsA} = k_{onA} [confA] + k_{offA}$$

$$k_{obsB} = k_{onB} [confB] + k_{offB}$$

Hemos reportado que las k_{off} de la unión de cada confórmer son cercanas y de magnitud despreciable frente a las magnitudes de k_{on} . Así, la velocidad de colisión depende de la k_{on} específica para cada confórmer y de la concentración de cada uno, en principio, desconocida. Como primera aproximación proponemos dos posibilidades por las que se observan las diferencias de k_{obs} : 1) la concentración de cada confórmer es la misma, por lo que las k_{on} han de ser distintas; o 2) las k_{on} son iguales y las concentraciones relativas de los confórmers son distintas.

Si la población relativa de las poblaciones conformacionales es igual, los efectos observados se deben a diferencias en las k_{on} intrínsecas de cada población. Dado que las macromoléculas son dímeros de la misma proteína, esto podría justificarse por un efecto diferencial de las cargas expuestas en cada confórmer sobre el campo electrostático resultante, lo

que influiría las velocidades de colisiones asistidas electrostáticamente. De cualquier manera, de existir una diferencia en las velocidades, esta no sería muy grande, dado que ambos eventos observados corresponden a colisiones bimoleculares limitadas por la difusión (Fig 3.1).

Por el contrario, si ambas poblaciones de conformeros tienen la misma k_{on} , las diferencias entre las velocidades observadas se deben a la población diferencial de cada *ensemble*. Si este fuera el caso, es posible estimar directamente la población relativa de cada uno, ya que:

$$k_{obsA} / k_{obsB} = [confA] / [confB]$$

Según esto, es posible estimar que las poblaciones relativas son 66% A, 33% B. Es decir, el 66% de la unión procedería por el camino de "dos estados". Para aclarar estos puntos profundizando en el estudio del mecanismo de unión, abordamos las interacciones de la interfase con métodos de ingeniería de proteínas, presentados en los siguientes capítulos.

ANÁLISIS DE LA INTERFASE PROTEÍNA-ADN POR MUTAGÉNESIS SITIODIRIGIDA

- Juanito, debes dibujar lo que ves.
- Yo veo lo que dibujo, maestro.

Juanito a su maestro de 2do grado, José María Firpo

ANÁLISIS DE LA INTERFASE PROTEÍNA-ADN POR MUTAGÉNESIS SITIODIRIGIDA

La distribución energética de todas las reacciones de unión entre macromoléculas es el resultado de intrincadas redes moleculares de contacto. La remoción puntual de los grupos interactuantes provee información sobre la contribución energética de cada uno de los contactos removidos, y de cómo éstos influyen sobre los demás componentes no modificados. Para una asignación racional de las energías asociadas es fundamental contar con estructuras de alta resolución de los componentes físicos en estudio, así como de una descripción detallada de la red de interacción, y de los aspectos dinámicos asociados a los mismos. En las condiciones actuales, es virtualmente imposible recabar todos los datos necesarios para interacciones tan complejas como el caso de las interacciones proteína-ADN, donde las macromoléculas pueden tener más de 2000 átomos, la superficie de interacción ser mayor a $1500 \text{ \AA}^2$, los contactos intermoleculares sumar decenas y los movimientos asociados, distribuirse por todo el sistema en forma correlacionada.

A pesar de la evidente complejidad, y con el fin de aportar a la comprensión de los mecanismos por los cuales las proteínas se unen específicamente a secuencias de ADN, el modelo de interacción de E2c-s35 provee un marco inicial plausible de ser analizado experimentalmente utilizando los métodos de medición descritos en los capítulos anteriores. Para este modelo, contamos además con la posibilidad de relacionar las mediciones energéticas con estructuras de alta resolución, tanto de los componentes libres como del complejo proteína-ADN. Así, planeamos disectar energéticamente la interacción E2c-ADN por medio de un estudio de mutagénesis racional sitio-dirigida.

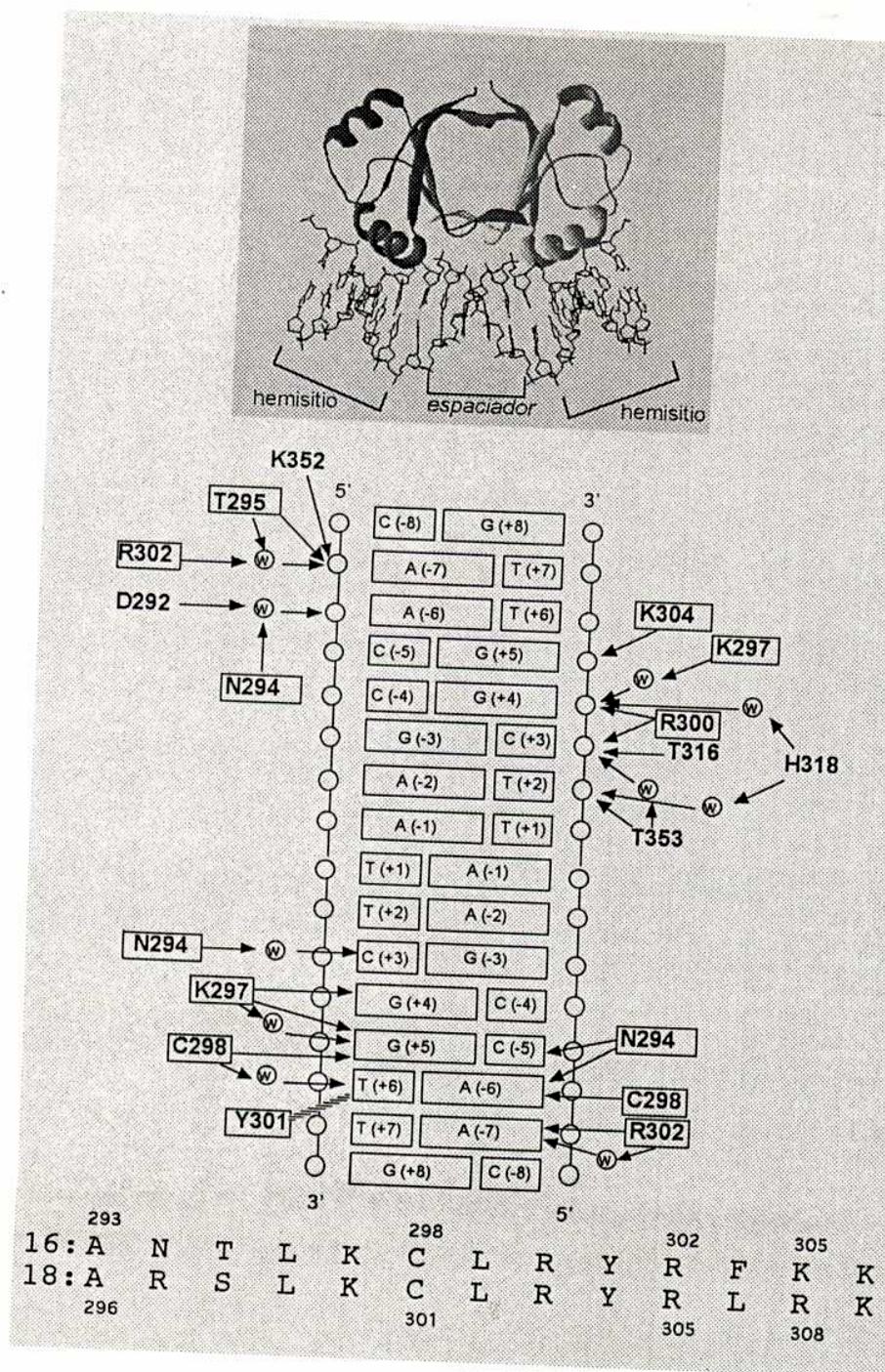
DISEÑO DE LAS MUTACIONES PUNTUALES

La estrategia general consistió en mapear energéticamente la superficie de interacción entre E2c y ADN por medio del reemplazo puntal de cada uno de los aminoácidos de E2c involucrados directamente en la interacción con ADN. Dado que no contamos con una estructura de alta resolución para el complejo (HPV-16) E2c-ADN, la identificación de los posibles blancos mutagénicos se realizó en base a las estructuras resueltas de los homólogos de E2c para los cuales existen estructuras del complejo con ADN: BPV-1 y HPV18. Estos modelos muestran que la interacción se da principalmente por las hélices mayores que se extienden por fuera del barril β -dimérico (también llamadas "hélices de reconocimiento" o hélice 1), y los surcos mayores del sitio canónico E2 en una conformación de B-ADN (Fig 4.1A). Dada la alta similitud de la secuencia primaria de E2c de los diferentes *papillomavirus* (que, en particular, es máxima en la zona de la hélice 1, Fig 4.1C) y la conservación estructural descrita para las estructuras resueltas de las diferentes cepas (incluida la HPV16, Fig 1.5), es de esperar que las zonas de contacto estén también altamente conservadas.

El análisis estructural de la interacción entre (HPV18) E2c y ADN específico revela una intrincada red de interacciones no-covalentes ^{34, 48} (Fig 4.1B), entre las que se destacan uniones electroestáticas, uniones de hidrógeno, interacciones de van der Waals, y "uniones hidrofóbicas". Se pueden agrupar dos categorías de contactos, los "inespecíficos" que ocurren entre E2c y los grupos fosfato de la cadena principal de ADN, y los "específicos" que se dan entre las bases nitrogenadas del sitio E2 y los residuos de los aminoácidos de la hélice 1 de E2c. Los "inespecíficos" contabilizan 24 uniones electroestáticas entre grupos cargados positivamente en E2c y los fosfatos de la cadena desoxirribonucleotídica. Dentro de los "específicos" se pueden contabilizar 14 uniones de hidrógeno directas, 8 uniones de hidrógeno mediadas por agua, y 2 interacciones de dispersión (Fig 4.1B).

Dada la alta conservación tanto del sitio E2 (ACCGnnnnCGGT), como de la topología de E2c, no es sorprendente encontrar que los aminoácidos más conservados son aquellos involucrados en los contactos directos entre E2c y las bases nitrogenadas del sitio E2: K297, C298, R300, R302, sugiriendo además que existe una conservación evolutiva en el mecanismo estructural de reconocimiento y discriminación de secuencia.

Figura 4.1: Esquemas estructurales de la interacción E2c-ADN específico. A) modelo cristalográfico derivado del co-cristal de HPV18-E2c:ADN. B) Mapa de los contactos entre E2c y ADN derivado de la estructura presentada en A. Para mayor claridad, en la parte superior se esquematizan los contactos con los grupos fosfato y en la inferior los contactos con las bases. Estos se repiten simétricamente sobre el centro del palíndromo. (w) es una unión de hidrógeno mediada por una molécula de agua. C) alineamiento de las secuencia primaria de la hélice de reconocimiento de E2c de HPV16 y HPV18.



DEFINICIÓN DE LAS REGIONES MUTAGENIZADAS

Región de la hélice 1: se mutagenizaron todos y cada uno de los aminoácidos de la hélice de reconocimiento, desde Ala293 hasta Lys305. La Fig4.2A muestra la superficie de E2c accesible al solvente, destacándose en la zona de la hélice 1 cada uno de los residuos modificados. Las mutaciones se realizaron de manera racional ²⁵, apuntando a la remoción de grupos funcionales de los diferentes aminoácidos, reemplazando de manera de causar la menor perturbación estructural posible. Por ejemplo, el reemplazo de Tyr por Phe o Ala permite asignar directamente los efectos al grupo OH (en el caso del reemplazo por Phe) o al anillo aromático (reemplazo por Ala). En algunos casos se buscó delecionar todos los grupos funcionales más allá del carbono- α , para lo cual el aminoácido fue reemplazado por Ala. En la tabla 4.1 se detallan las mutaciones realizadas sobre los residuos mostrados en la Fig 4.2B.

La superficie total de la hélice1 de E2c involucrada en la interacción con ADN es de ~ 900 Å². Dado que mutagenizamos la totalidad de los residuos expuestos de dicha región, podemos decir con confianza que el mapeo energético cubre más del 95% de la superficie total de interacción.

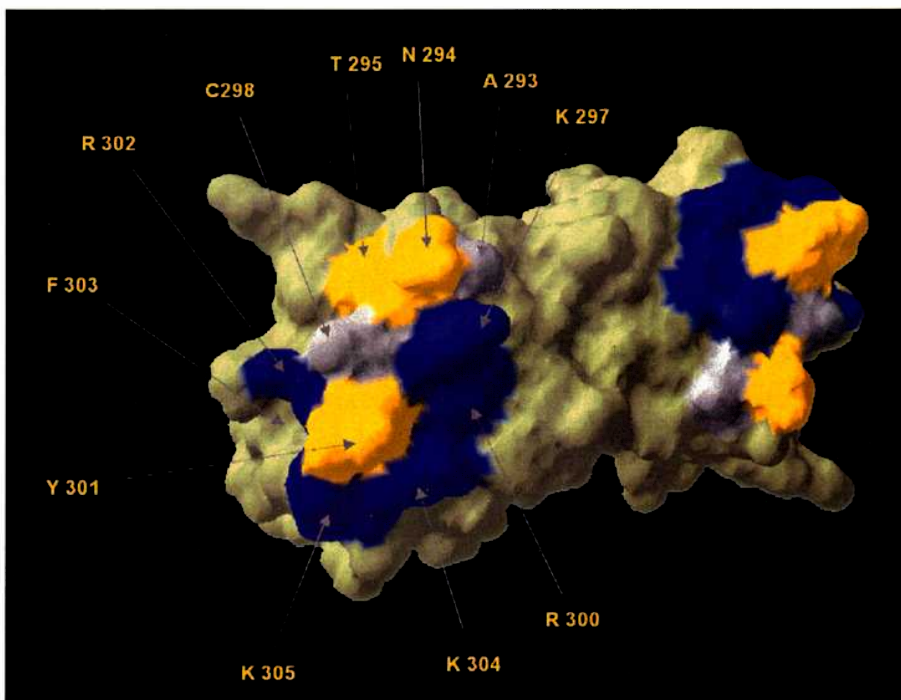
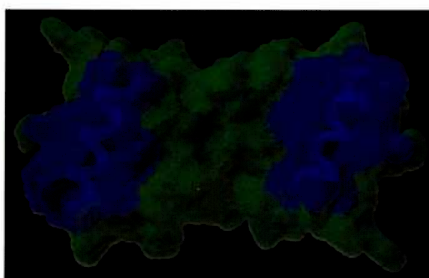
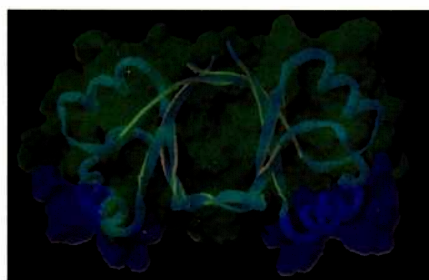
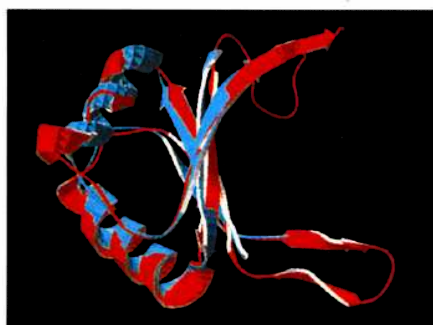


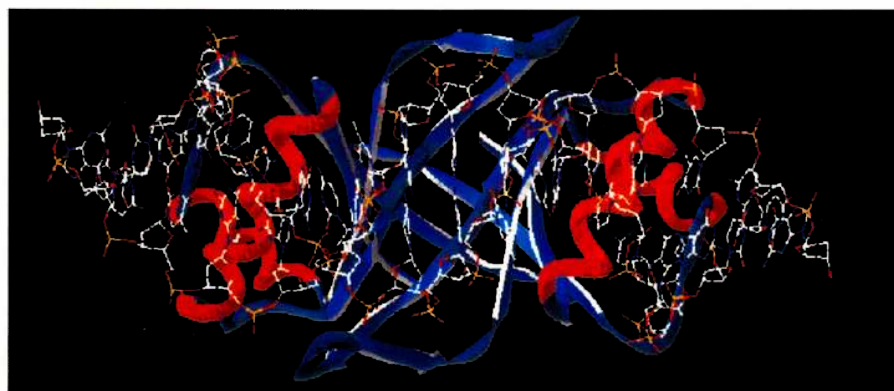
Figura 4.2: mutaciones realizadas sobre la región de la hélice 1. A) superficie de E2c accesible al solvente, según la estructura cristalográfica de E2c de HPV16 (*). La superficie correspondiente a la hélice 1 se muestra en azul. B) Representación individual de los aminoácidos mutagenizados. Los colores corresponden a convenciones internacionales: amarillo, neutro polar; en blanco, hidrofóbico; en azul, básico.

Región del loop $\beta 2 - \beta 3$: en los primeros estudios estructurales de la interacción BPV1-ADN, se identificó otra región de contacto con ADN más allá de la hélice1³⁴. Se trata de un *loop* que conecta la $\beta 2$ y la $\beta 3$ del barril β -dimérico, para el cual se identificaron contactos directos a fosfatos de la cadena nucleotídica, uniendo la zona de la secuencia "separadora" de los hemisitos E2 del ADN (Fig 4.3). Posteriormente se encontró que este *loop* no está definido en la densidad electrónica de ningún cristal de E2c libre, y se propuso que esta región presentaría una alta movilidad cuando la proteína está en estado "libre", ordenándose al contactar ADN. Sin embargo, en el complejo de HPV18-E2c:ADN, no se observó densidad electrónica definida para esta región, por lo que se argumentó que, para E2c de HPV, este *loop* debería mantener la movilidad incluso después de unido a ADN^{36, 37, 48}. En base a esto, se propuso que los contactos definidos para el *loop* solo se dan en las cepas de BPV1 y no de HPV, justificando en forma *ad-hoc* la diferencia de discriminación que ambas cepas presentan cuando se ensaya la unión con sitios E2 mutantes en la secuencia "separadora"³⁹.

Con el fin de aclarar este punto y medir la contribución de la región del *loop* $\beta 2 - \beta 3$ al mecanismo de unión a ADN de la cepa de HPV16, planeamos 5 mutaciones puntuales sobre 3 aminoácidos del *loop* (tabla 4.1)



loop $\beta 2 - \beta 3$



hemisito separador hemisito

Figura 4.3: *loop* $\beta 2 - \beta 3$. A) modelo cristalográfico derivado del co-cristal de BPV1-E2c:ADN. B) superposición de las estructuras de E2c derivadas de los cristales de BPV1-E2c:ADN, rojo y HPV16-E2c, verde. Se muestra solo un monómero.

CONSTRUCCIÓN DEL CONJUNTO DE MUTANTES

La descripción detallada de la construcción de conjunto de mutantes se presenta en Materiales y Métodos (capítulo 7). Brevemente: se diseñaron cebadores conteniendo la mutación deseada y, utilizando como molde un plásmido de expresión conteniendo el marco de lectura de E2c, se amplificó por PCR todo el plásmido. Secuenciando el ADN recuperado se confirmó la introducción de la mutación deseada y la integridad de resto del marco de lectura de E2c. Las proteínas fueron expresadas en forma recombinante en *E. coli* y la purificación se siguió de forma similar a la descrita para la proteína salvaje.

Salvando engorrosas dificultades técnicas surgidas, pudimos construir el total de las mutaciones planeadas. Exceptuando dos mutantes, la expresión recombinante fue exitosa y obtuvimos proteína soluble en todos los casos. Para las mutaciones L296A y L299A, no se observó banda de sobreexpresión en las condiciones estándar del ensayo. Este resultado no es del todo sorprendente, ya que ambas cadenas laterales mapean en la cara interna de la anfipática hélice 1, donde existe un "core" hidrofóbico formado con la cara externa del barril- β . Es posible que este "core" esté involucrado en la estabilización de la hélice1, el cual puede ser perturbado por las mutaciones Leu a Ala, desestabilizando el estado plegado de la proteína toda. Si esto ocurriera, no es de extrañar la falta de expresión consistente de las mutantes antedichas. De todas formas esto no invalida el estudio mutagénico, ya que estos residuos son predichos como no-contactados por ADN.

Las mutantes restantes fueron purificadas a homogeneidad, obteniéndose un rendimiento entre 2 y 40 mg de proteína por litro de cultivo. Cabe destacar que el último paso de purificación utilizado es una columna de filtración por tamiz molecular, donde todas las proteínas eluyeron en el volumen esperado para un dímero de E2c (~19 KDa), indicando que las mutaciones introducidas no afectan la dimerización de E2c.

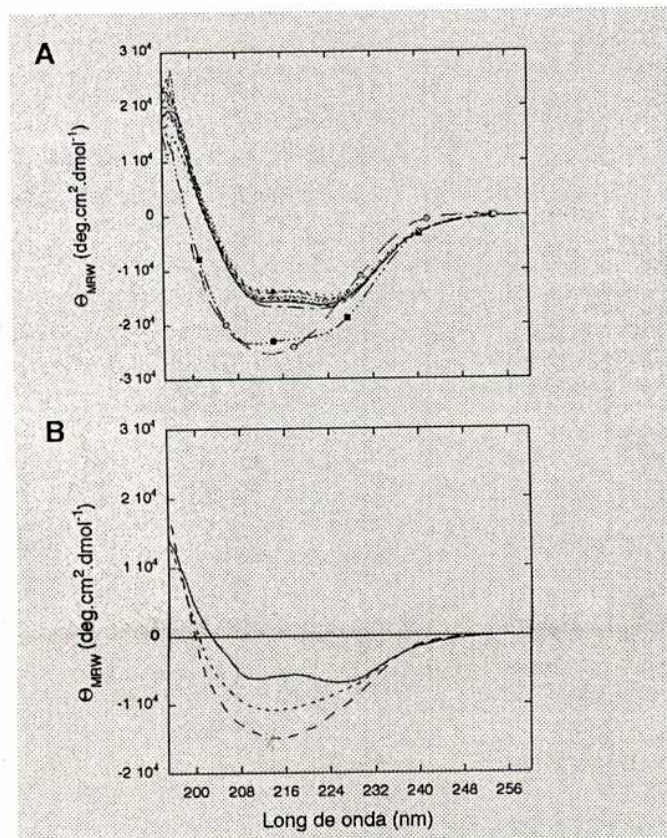
A juzgar por los espectros de dicroísmo circular (DC) en el UV lejano, todas las proteínas mutantes purificadas se encuentran plegadas (Fig 4.4). En general, el estado de plegamiento, en tanto a topología general y contenido de elementos de estructura secundaria, no se encuentra afectado drásticamente por la mutación. Sin embargo, hay dos excepciones que destacar, situadas ambas en el N-cap de la hélice1:

La mutación A293G presenta un espectro de DC anómalo, con una disminución de las bandas características de α -hélice, y un aumento de la intensidad negativa alrededor de 215-217 nm. En esta zona del espectro contribuyen principalmente las láminas- β y otros motivos estructurales como *loops* y *hairpins*, por tanto el cambio dicroico no es unívocamente asignable a un rearrreglo estructural definido ³⁰. Dado que se trata de una mutación puntual, asumimos que la aparición de señales en la región de 215-217nm se da a expensas de la desaparición de las señales de α -hélice. Esta hipótesis se refuerza si consideramos que el cambio introducido se localiza sobre el N-cap de la hélice1, siendo además un reemplazo por Gly, aminoácido que es considerado como "desestabilizador" de conformaciones α -hélice ²⁵. Estos resultados sugieren que la mutación introducida provoca una desestabilización local de la hélice 1 de E2c, favoreciendo un plegamiento alternativo de la cadena polipeptídica.

Este estado de plegamiento alternativo del dímero de E2c puede también evidenciarse en la proteína salvaje: si E2c es calentada a 50°C por 15 min, temperatura para la cual el plegamiento general de E2c no se afecta, el espectro de DC recuperado presenta las mismas características que el descrito para A293G (Fig 4.4B). Como se mencionó en el capítulo 1, la hélice 1 de E2c no es una región con estructura fuertemente consolidada, e incluso puede considerarse como un "foldón" o microdominio de plegamiento pseudo-independiente del barril- β ^{54, 72, 73} (Nadra *et al.* enviado). Dada la evidente desestabilización causada por la delección de un metilo en el N-cap de la hélice, podemos atribuir el efecto al aumento de grados de libertad de los ángulos phi y psi en la posición 293, efecto típico atribuido a la mutación por Gly.

En cuanto a la mutación en la posición 294 (Asn-> Ala), el espectro presenta un aumento en la elipticidad absoluta de las bandas características de α -hélice, al mismo tiempo que un corrimiento del mínimo de 211 hacia 208 nm. Esto sugiere que la mutación introducida consolida la formación de la hélice1, a expensas de algún elemento de estructura secundaria que contribuye en la región de 212-218 nm.

Figura 4.4: Espectros de dicroísmo circular de las mutantes de E2c. A) Se colectaron los espectros de DC en el UV lejano de cada una de las proteínas purificadas en 25mM AcNa pH=5.6, 10mM NaCl, 10mM MgCl₂, 1mM DTT a 25°C. Se destacan E2c salvaje (línea llena), A293G (puntos blancos) y N294A (cuadrados negros) B) Espectro de la proteína salvaje en las mismas condiciones a 25°C (línea llena), a 50°C (línea punteada), y a 25°C luego del tratamiento térmico (líneas largas).



DISECCIÓN DE LA ENERGÍA LIBRE DE UNIÓN

Utilizando el ensayo de unión a ADN desarrollado (capítulo 2), y contando con reactivos de alto grado de pureza, realizamos mediciones de unión entre un oligonucleótido conteniendo un sitio de unión específico de E2 (s35) y cada una de las mutantes construidas. Exceptuando la mutante A293G, todas las proteínas formaron complejos solubles y reversibles con ADN. En general, los ensayos se realizaron a baja concentración a fin de minimizar posibles efectos de uniones con estequiometrías complejas al sitio 35, como lo descrito para la proteína salvaje (capítulo 2). En este sentido, todas las Isotermas de unión pudieron ser bien descritas con ajustes a modelos simples de estequiometría 1:1 (ecuación 7.1, capítulo 7). Realizamos entre 3 y 5 Isotermas de unión para cada una de las proteínas purificadas, incluyendo a la proteína salvaje, en las mismas condiciones de presión, temperatura y composición del tampón.

De los ajustes de los datos calculamos el valor de la constante de disociación (K_D) y, teniendo en cuenta el valor de referencia para la proteína salvaje, calculamos el efecto de la mutación sobre la energía libre de Gibbs según:

$$\Delta\Delta G = \Delta G_{\text{mut}} - \Delta G_{\text{wt}}$$

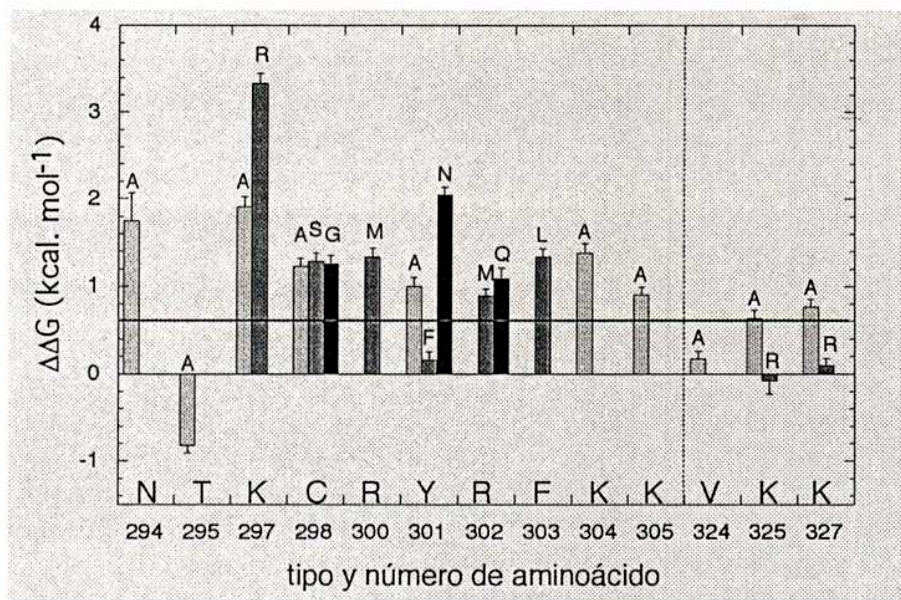
$$\Delta\Delta G = -RT \ln (K_{D\text{mut}} / K_{D\text{wt}})$$

Los resultados se muestran en la tabla 4.1 y en la figura 4.5.

Tabla 4.1: Mutaciones puntuales realizadas sobre E2c: constantes de disociación de la unión a FI-s35 y diferencia de energía libre de Gibbs.

Posición	Mutación	K_D (nM)	$\Delta\Delta G$ (kcal.mol ⁻¹)
wt		1.20 ± 0.16	0
A293	G	ND	NA
N294	D	NM	NA
	A	23.1 ± 11	1.75 ± 0.3
T295	A	0.40 ± 0.05	-0.82 ± 0.12
K297	A	29.4 ± 5	1.89 ± 0.12
	R	336 ± 54	3.33 ± 0.12
C298	S	10.1 ± 1	1.25 ± 0.1
	G	10.5 ± 1.2	1.28 ± 0.1
	A	9.5 ± 1	1.22 ± 0.1
R300	M	11.4 ± 1.4	1.33 ± 0.1
Y301	N	38.0 ± 2	2.04 ± 0.08
	F	1.58 ± 0.2	0.16 ± 0.09
	A	6.5 ± 0.7	1.00 ± 0.1
R302	Q	7.5 ± 1.3	1.08 ± 0.13
	M	5.4 ± 0.3	0.89 ± 0.08
F303	L	11.4 ± 1.1	1.33 ± 0.1
K304	A	12.4 ± 1.5	1.38 ± 0.1
K305	A	5.5 ± 0.3	0.90 ± 0.08
V324	A	1.6 ± 0.1	0.17 ± 0.09
K325	A	2.5 ± 0.2	0.52 ± 0.09
	R	1.04 ± 0.3	-0.08 ± 0.15
K327	A	4.3 ± 0.2	0.75 ± 0.08
	R	1.4 ± 0.1	0.091 ± 0.09

Figura 4.5: Efecto de las mutaciones puntuales en la unión de ADN específico. Se realizaron entre 3 y 5 isotermas de unión entre cada proteína y el oligonucleótido FI-s35. Se muestran los valores de la diferencia de energía libre de Gibbs respecto a la proteína salvaje, listados en la tabla 4.1. La línea punteada corresponde al valor de la energía térmica a 25°C (0,6 kcal/mol).



Para el análisis de los resultados supondremos que el reemplazo afecta sólo a la interacción entre el residuo mutado y el ligando. Consideramos que las perturbaciones realizadas son mínimas y, como indican los espectros de DC, no afectan la topología de plegamiento. Descartaremos por tanto la posibilidad de que existan cambios en la estabilidad de otros enlaces de la red de interacciones que pudieran compensar los efectos de la mutación; vale decir, supondremos que los efectos de las mutaciones son independientes.

Exceptuando a la mutación T295A, todas las sustituciones causan una disminución en la afinidad de E2c por el oligonucleótido FI-s35. Esto está en perfecta relación con lo esperado, pues puede interpretarse que la mutación por Ala causa un efecto análogo a la delección de los contactos específicos mediados por el residuo más allá del carbono-α. De esta manera, la energía asociada al evento de unión disminuye proporcionalmente con la energía de los enlaces deleccionados, validando el abordaje utilizado.

Lo primero que se pone en evidencia es que los valores absolutos de las diferencias energéticas medidas son pequeños, entre 0 y 2 kcal/mol. El valor de referencia usual es el de la energía térmica a 25°C ($RT = 0,6$ kcal/mol). De todas maneras, el error acumulado en las mediciones es cercano al 10%, indicando que el ensayo de unión desarrollado es lo suficientemente sensible para analizar los pequeños efectos de las mutaciones sobre la energía de unión. Los valores de $\Delta\Delta G$ obtenidos para las sustituciones puntuales están en el rango de los medidos para otros sistemas, por otros métodos, tanto en estudios de plegamiento de proteínas²⁵, interacción proteína-proteína^{8, 109}, e interacción proteína-ADN^{24, 47, 63}.

AUSENCIA DE "HOT-SPOTS" EN LA INTERFASE DE UNIÓN

En los últimos años se han acumulado gran cantidad de estudios sobre interfaces de unión proteína-ligando, particularmente los detallados por métodos de mutagénesis extensiva sobre interfases ¹⁸. Una característica general emergente de estos estudios es la ocurrencia de residuos *calientes*, o "hot-spots" en las interfases. Estos se definen como aquellos sitios para los cuales el reemplazo de un grupo interactuante produce una disminución desproporcionada de la energía de asociación, respecto al efecto del reemplazo de otros residuos de la misma interfase. Los valores energéticos asociados a la modificación de un "hot-spot" son generalmente mayores a 3 kcal/mol, siendo así caracterizados como contactos *claves* en las interfases, en tanto que existe algún tipo de cooperatividad en la Interacción pivotando alrededor de estos residuos ¹⁸.

Dada la extensiva mutagénesis realizada, que cubre más del 95% de la superficie de interacción de E2c con ADN, y la medición de las diferencias energéticas presentadas (Fig 4.5), podemos decir con confianza que no existen "hot-spots" en la interfase de interacción E2c-s35.

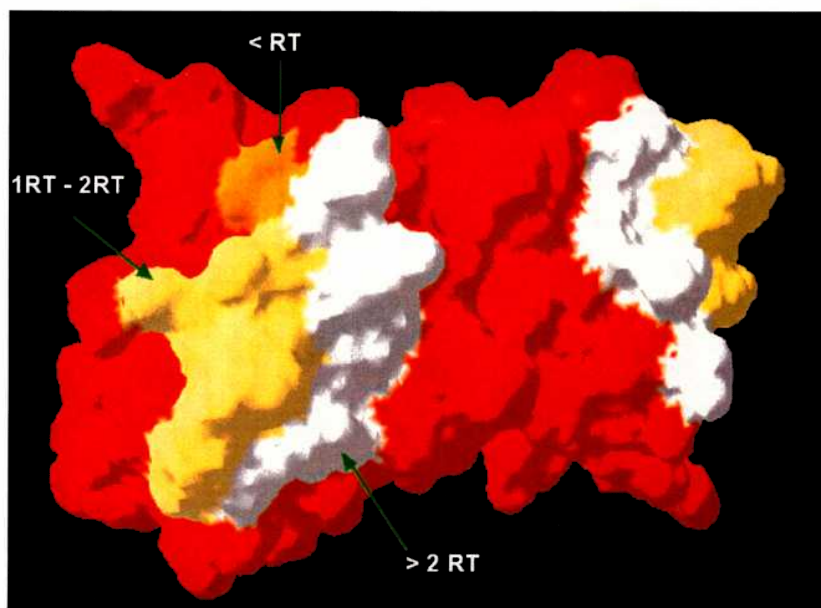
Estudios estadísticos de los atributos estructurales de los "hot-spots" muestran que éstos se encuentran preferentemente en sitios deshidratados de las interfases ⁸. En este sentido, la ausencia de hot-spots en la interacción E2c-ADN puede ser asociado a una interfase parcialmente hidratada, en concordancia con los estudios cristalográficos del complejo, donde se describen densidades electrónicas correspondientes a moléculas de agua mediando interacciones (Fig 4.1C).

Una forma alternativa de interpretar la caracterización energética presentada (Fig 4.5) es que la energía libre de unión está *delocalizada* por el sistema todo, en el sentido que no existen contactos críticos mediando la interfase. Consideraciones teóricas recientes muestran que una delocalización energética de este tipo es evidencia indirecta de la ocurrencia de un proceso de plegamiento asociado al evento de unión ^{23, 24}. El acoplamiento entre plegamiento y unión ha sido descrito en detalle para las interacciones proteína-ADN en general ²⁵ y en particular para el sistema E2c-ADN (capítulo 3, ^{26, 26}). Por tanto, el modelo de interacción E2c-ADN es tal vez el primer caso en el cual se comprobó experimentalmente una correlación entre el plegamiento asociado a la unión de ligando y la ausencia de "hot-spots".

LOCALIZACIÓN ENERGÉTICA EN LA ESTRUCTURA PROTEICA

A fin de disectar los efectos de las mutaciones, categorizamos los residuos en tres grupos. Estos fueron agrupados respecto al valor de $\Delta\Delta G$ tomando como referencia el valor de la energía térmica a 25°C (Fig 4.5). La representación coloreada de las categorías sobre la estructura cristalográfica se muestra en la Fig 4.6.

Figura 4.6: Localización estructural de las contribuciones energéticas. Se muestra la superficie accesible al solvente del dímero de E2c. La zona correspondiente a las "hélices 1" fueron coloreada según la perturbación energética que causan las mutaciones puntuales sobre la unión de ADN específico. Se tomaron los valores correspondientes a las mutaciones a Ala (o Met), mostrados en la figura 4.5.



El patrón de distribución de las contribuciones energéticas de cada residuo parece no ser aleatorio. Aquellas que contribuyen en más de 2 RT se encuentran agrupadas en la cara interna de la hélice de unión, respecto al eje de simetría del homodímero. Dada la estructura de E2c unido a ADN, los residuos que se encuentran más cercanos al centro del palíndromo presentan una mayor contribución energética que aquellos expuestos hacia los extremos. Para interacciones proteína-proteína, los patrones de distribución energética suelen estar correlacionados con parámetros estructurales tales como la superficie accesible al solvente o la superficie molecular^{8, 25}. Para el caso de E2c-ADN, no encontramos correlación alguna entre el patrón energético y la superficie total del residuo mutado, como tampoco con la diferencia en superficie accesible al solvente causada por la mutación (no mostrado). Tampoco encontramos una correlación entre las diferencias energéticas medidas con el número de contactos que cada aminoácido sustituido realiza, como se presenta en las secciones subsiguientes.

ADITIVIDAD DE LAS ENERGÍAS ASOCIADAS A LOS RESIDUOS INDIVIDUALES

Para describir el efecto causado por las mutaciones puntuales, supusimos que la introducción de la mutación afecta solo al residuo mutado. De esta manera descartamos la eventual ocurrencia de efectos cooperativos o compensaciones energéticas distantes al sitio de mutación. En este sentido, la ausencia de "hot-spots" apoya indirectamente este supuesto. Por otro lado, si los efectos causados por las mutaciones son independientes entre sí, las energías asociadas a cada una de las sustituciones deben ser aditivas ²¹. El mapeo energético presentado incluye toda la hélice de reconocimiento y 3 residuos del *loop* $\beta 2$ - $\beta 3$ y si, como asumimos, los efectos de las mutantes fueran independientes, es de esperar que la sumatoria de las energías asignadas a cada residuo sea igual a la energía total asignada a la interacción de la proteína salvaje y el sitio de ADN. Esta última puede ser calculada a partir de la constante de disociación del equilibrio E2c-s35.

La sumatoria de las energías libres asociadas a las mutaciones de delección total de grupo (Ala o Met) es de 11.5 ± 0.7 kcal/mol, en tanto que la energía libre total de la interacción E2c-ADN es de 11.9 ± 0.1 kcal/mol, medida en las mismas condiciones de entorno. Esto indica que las perturbaciones de las interacciones entre los grupos funcionales de los residuos y las bases son independientes entre sí, validando la asunción original de no-cooperatividad.

Buscando una confirmación directa de la independencia del efecto de las mutaciones al mecanismo de unión de E2c-s35 al equilibrio, utilizamos el abordaje de ciclos de doble mutación (capítulo 7). Para ello construimos, expresamos y purificamos tres E2c en las cuales introducimos 2 mutaciones a Ala simultáneamente. Para cada una de ellas realizamos isotermas de unión y calculamos la diferencia en la energía libre de unión al sitio E2, como en los casos anteriores. Si las mutaciones fueran independientes, el efecto energético sobre la doble mutante ha de ser igual a la sumatoria de las energías asociadas a cada una de las mutaciones consideradas en forma separada. Los resultados presentados en la tabla 4.2 indican que las mutaciones se comportan dentro del rango esperado, es decir, son independientes.

Tabla 4.2: aditividad de las energías libres de unión a ADN ensayada por sustituciones dobles

	C298A N294A	C298A R300M
$\Delta\Delta G$ predicho ^a	$2,5 \pm 0,2$	$2,9 \pm 0,3$
$\Delta\Delta G$ medido	$3,1 \pm 0,6$	$2,7 \pm 0,4$

EFFECTO DE LAS MUTACIONES SOBRE LA UNIÓN DE ADN INESPECÍFICO

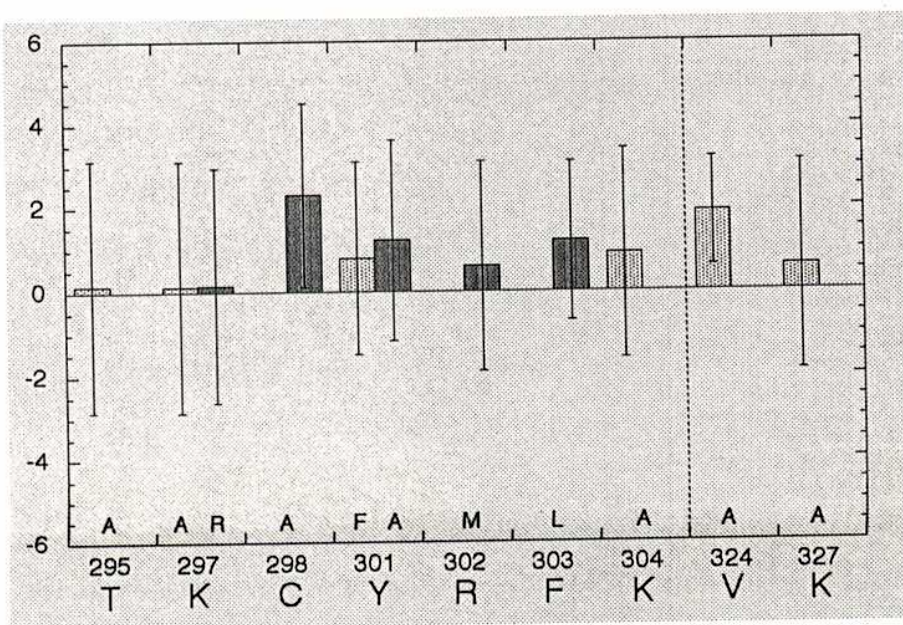
Para evaluar el efecto causado por las mutaciones en el reconocimiento inespecífico de ADN utilizamos el oligonucleótido iset35, que contiene la misma composición de bases que el site35, pero distribuidas al azar. Los resultados obtenidos para el análisis de un grupo de mutantes se muestran en la Fig 4.7

Dada la baja afinidad de la interacción E2c - ADN inespecífico, el método de medición utilizado presenta altos márgenes de error. Sin embargo, es posible estimar que la diferencia energética presentada por las mutantes respecto a E2c salvaje es mínima o nula. Es decir, las mutaciones introducidas no causan efectos significativos en la unión inespecífica de ADN. Dadas las características generales descritas para la unión inespecífica de ADN (ver capítulo 1), es esperable que una perturbación pequeña, como las mutaciones puntuales aquí realizadas, no alteren sustancialmente la interacción inespecífica.

Si bien no es posible afirmarlo con certeza debido al error asociado, el efecto de aditividad de las energías libres descrito para la unión específica parece no cumplirse para la unión inespecífica de ADN, en concordancia con lo encontrado en otros sistemas de interacción proteína-ADN ⁷⁵.

Figura 4.7: Efecto de las mutaciones puntuales en la unión de ADN inespecífico.

Se realizaron entre 2 y 3 isotermas de unión entre cada proteína y el oligonucleótido FI-iset35. Se muestran los valores de la diferencia energía libre de Gibbs respecto a la proteína salvaje.



ANÁLISIS DE ALGUNOS CONTACTOS PUNTUALES, PARTICULARMENTE DE AQUELLOS EN LOS CUALES SE LLEVÓ A CABO MÁS DE UN REEMPLAZO AMINOACÍDICO

El método de mutación y análisis desarrollado es lo suficientemente sensible para detectar pequeñas contribuciones energéticas. Si se analiza más de una mutación puntual realizada sobre el mismo residuo, debería ser posible asignar contribuciones energéticas a los diferentes grupos funcionales de los aminoácidos de E2c. En nuestro plan de mutagénesis racional incluimos la construcción y medición de más de una mutación para 4 aminoácidos de la hélice 1, y de 2 del *loop* β 2- β 3. Analizaremos los resultados para cada uno de esos reemplazos, relacionando las estructuras resueltas y la termodinámica de interacción.

LISINAS 304 Y 305

Estos residuos se sitúan en el extremo C-terminal de la hélice de reconocimiento de E2c. La Lys 304 (Arg en la estructura resuelta para HPV18:ADN) ha sido descrita formando sólo una unión electrostática con los grupos fosfato de G(4), mientras que a Lys305 no se le asignaron interacciones con ADN. Tanto la sustitución de una u otra causan una pérdida de afinidad entre 0,9 y 1,5 kcal/mol (Fig 4.5), valores en concordancia con lo descrito para uniones a fosfatos en el sistema reconocimiento de BamHI y ADN específico, medidos por otros métodos^{24, 25}.

Como la sustitución de K305 causa una perturbación en la unión, sugerimos que este residuo interactúa con ADN, presumiblemente con grupos fosfato vecinos. Dados los valores energéticos y los errores asociados, es plausible asignar a K305 una unión del mismo tipo que la que asignada a K304. El hecho que esto no se observe en la estructura cristalográfica sugiere que, o bien los contactos varían entre la cepa de HPV16 y 18 o, lo que es más probable, la estructura cristalográfica obtenida no permite resolver interacciones de estas características, dado que solamente se miden distancias medias de las disposiciones espaciales de los componentes en la red cristalina.

ARGININA 300 Y TREONINA 295

Conjuntamente con Lys 304, éstos son los únicos aminoácidos asignados solamente a interacciones con grupos fosfato. Cabe mencionar que para el homólogo de HPV18, en el cual esta basada la asignación estructural de la interacción con ADN, hay una Ser en la posición 295 (Fig 4.1). El reemplazo de Thr 295 por Ala resulta en una proteína más afín por $\approx$ 35 que la proteína salvaje. Este comportamiento, en principio inesperado, puede ser adjudicado a la presencia de interacciones "desfavorables" formadas por T295, o por cambios en las energías de interacción de otros residuos no modificados. Hasta no realizar una sustitución doble sobre T295 y otros residuos, no nos es posible adjudicar unívocamente el efecto observado.

TIROSINA 301

¿

La sustitución de Arg 300 por Met causa una pérdida de alrededor de 1,3 kcal/mol de la energía libre de Gibbs de la unión a s35. Este efecto puede ser directamente asignado a la pérdida de la unión salina a los fosfatos del ADN, ya que las posibles contribuciones hidrofóbicas de Arg y Met son comparables. Valores energéticos comparables han sido observados para la asignación de las uniones fosfato de la interacción entre BamHI y ADN específico ^{24, 27}.

Según las estructuras resueltas de los co-cristales de E2c-DNA, existe un residuo de Tyr en la zona de contacto de la hélice 1. Este grupo estaría involucrado en un contacto dispersivo de apilamiento (*stacking*) con los anillos nitrogenados de las bases T(6) del sitio E2, descartándose interacciones del grupo OH de la Tyr (Fig 4.8A). Los resultados obtenidos para las diferentes sustituciones realizadas sobre dicho residuo se muestran en la tabla 4.1 y la Fig 4.5.

El reemplazo por Phe no afecta la interacción con site35, indicando que el grupo OH no está involucrado en las interacciones medidas al equilibrio, en absoluta concordancia con los resultados cristalográficos descriptos para los homólogos de HPV18 y BPV1. Esto sugiere además que el contacto asignable a Tyr301 sería análogo en HPV16. La delección completa del grupo fenol, resultante de la mutación por Ala, afecta en alrededor de 1 kcal/mol la energía libre de unión. Descartando posibles efectos compensatorios, esto sugiere que la interacción de *stacking* descrita contribuye en valores cercanos a 0,5 kcal/mol por residuo, por monómero, en concordancia con las estimaciones realizadas en otros sistemas ^{28, 29}.

Varias cepas de *papillomavirus* poseen un residuo de Asn en la posición 301, por lo que ensayamos una sustitución de este tipo en la cepa de referencia, HPV16. Los resultados muestran que, sumado al efecto de la delección del grupo fenilo, la introducción del grupo amida afecta considerablemente la energía de interacción. Esto sugiere que la sustitución del anillo aromático por un grupo polar afecta las contribuciones hidrofóbicas de la unión, y que éstas no son reemplazadas por otras interacciones favorables.

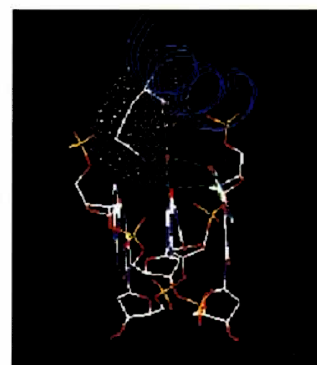
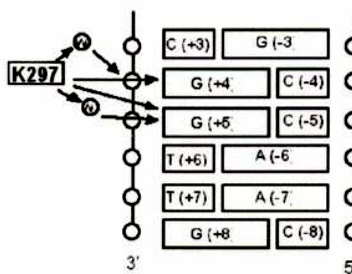
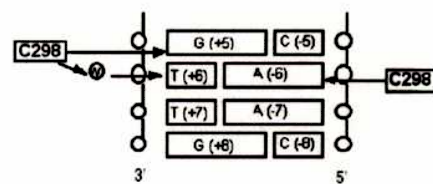
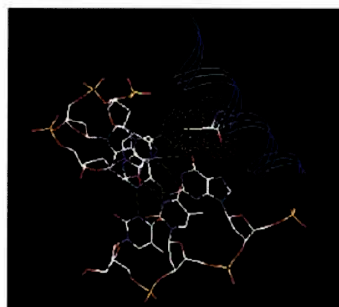
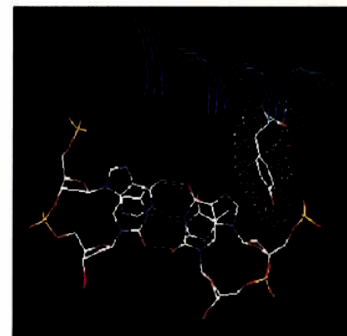
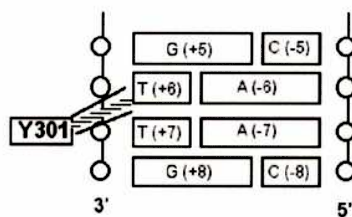


Figura 4.8: Contactos directo aminoácido-ADN. Se esquematizan los contactos asignados a tres de los aminoácidos mutagenizados de la hélice 1 de E2c, según la estructura resuelta del co-cristal de HPV18-ADN. A) Tyr301, B) Cys298, C) Lys297

CISTEÍNA 298

La presencia de un residuo de Cys en contacto directo con ADN es un fenómeno poco frecuente ⁴². Cuando se describió la estructura del primer complejo (BPV1) E2c:ADN, se sugirió que la Cys 298 estaría involucrada en la formación de 2 uniones de hidrógeno específicas con grupos de las bases G(5) y T(6) ³⁴. Para la estructura del complejo de (HPV18)E2c:ADN, se modelaron también interacciones directas con las bases, además de una unión de hidrógeno mediada por agua (Fig 4.8B).

Todas las sustituciones realizadas sobre el residuo 298, tanto a Ala como a Gly o Ser muestran una disminución en la energía de unión de alrededor de 1,2 kcal/mol. Dadas las diferentes características de estos grupos, es poco probable que la igualdad en el valor medido refleje que los mismos efectos estén teniendo lugar para cada reemplazo. Por ejemplo, se ha mostrado que el grupo OH de la Ser es capaz de actuar como dador de unión de H y, en algunos casos, reemplazar la Cys ²⁵. El hecho que tanto un reemplazo de Cys por Ser como por Ala muestren el mismo valor de perturbación total indica que, o bien Ser no es capaz de formar las uniones de H esperadas, tal vez por restricciones conformacionales, o que las que forma no son energéticamente equivalentes. Este no es un hecho del todo sorprendente, ya que es sabido que la energía de las interacciones no covalentes depende muy fuertemente de las características químicas y de las distancias interatómicas de los grupos interactuantes ⁴⁴. Por ejemplo, el valor energético de un enlace de hidrógeno puede variar considerablemente (más de 2 órdenes de magnitud) según los átomos dadores y aceptores, así como de los parámetros geométricos asociados al enlace ⁴⁴. Por otro lado debemos considerar que el átomo de azufre tiene un fuerte carácter hidrofóbico que, dado el microentorno particular de la interfase proteína-ADN, puede contribuir significativamente en la interacción.

El hecho que la sustitución de Cys tanto por Ala o Gly afecten a la energía libre de la unión en el mismo rango, sugiere que no hay restricciones conformacionales sobre la cadena polipeptídica principal mediadas por el residuo de la posición 298. El reemplazo de cualquiera de éstos por un átomo de hidrógeno, lo que otorga más grados de libertad alrededor del enlace peptídico, no afecta diferencialmente la energía de interacción total.

LISINA 297

La presencia de Lys en sitios de unión a ADN está ampliamente documentada ⁴². Se la encuentra involucrada tanto en uniones salinas con grupos fosfato, como en enlaces de hidrógeno simples o múltiples con las bases nitrogenadas, participando en contactos específicos e inespecíficos. Para el caso de la Lys297 de E2c, se propone que participa en 3 contactos con las bases G(4) y G(5), dos directos y uno mediado por agua, además de un contacto mediado por agua con un grupo fosfato vecino (Fig 4.8C).

La delección del residuo más allá del carbono- α causa la mayor pérdida de energía libre de todos los reemplazos por Ala efectuados sobre E2c (tabla 4.1, Fig 4.5), compatible con

las múltiples interacciones asignadas a este residuo. El reemplazo de Ala por Arg afecta aún más la interacción con ADN, mostrando que no sólo el carácter básico del aminoácido es responsable de las interacciones en esta posición. Más aún, el hecho que la sustitución resulte desfavorable en 3,4 kcal/mol, es indicativo de que otros efectos pueden estar teniendo lugar, como por ejemplo algún tipo de impedimento estérico dado el volumen del grupo guanidino de la Arg. Como se detalla a continuación, no siempre es posible observar estos efectos.

LISINAS 325 Y 327

Estos residuos se sitúan en el *loop* $\beta 2$ - $\beta 3$, para el cual no hay estructuras resueltas para ninguna cepa de HPV. En el caso de BVP1, se describieron contactos salinos entre aminoácidos básicos del *loop* y grupos fosfatos del "separador" de los hemisitios E2 del ADN.

El reemplazo por Ala causa una pérdida de alrededor de 1 kcal/mol, en el rango obtenido por la sustitución a Ala de la Lys 305 de la hélice de unión, asignada a un contacto salino débil. En concordancia con la hipótesis de contribución electrostática, la restitución de la carga puntual, por el reemplazo a Arg, resulta en el reestablecimiento de la energía libre de unión a los valores salvajes (tabla 4.1, Fig 4.5). Comparando estos datos con los presentados para la Lys297 y Lys304, conjuntamente con las referencias estructurales, es posible asignar al *loop* $\beta 2$ - $\beta 3$ como una región de alta movilidad, involucrada en contactos débiles y por tanto transientes con uno o más grupos fosfato del ADN. A diferencia de lo que ocurre con los contactos íntimos de la hélice, esta región es capaz de "acomodar" perturbaciones estructurales si la contribución electrostática del residuo se conserva.

ALANINA 293

Todos los ensayos de unión de E2c (A293G) con ADN resultaron en la formación de complejos proteína:ADN insolubles. Los agregados formados hacen que la unión a ADN sea irreversible, lo que impide una estimación de la energía libre de unión. Esta mutante presenta un espectro de DC anómalo (Fig 4.4), sin evidencias de elementos de hélice- α , atribuido a un plegamiento alternativo de la cadena polipeptídica. Como se mencionó, un espectro de las mismas características puede ser recuperado para la proteína salvaje como resultante de un tratamiento térmico suave. En concordancia, los ensayos de unión realizados con una muestra salvaje de estas características resultan en la formación de agregados insolubles proteína:ADN (no mostrado). Esto sugiere que este plegamiento alternativo del polipéptido E2c puede interactuar con ADN, pero el complejo formado posee características estructurales diferentes a los anteriormente descritos, tal que resultan en múltiples interacciones que promueven agregación macromolecular. Estudios paralelos llevados a cabo en otras líneas del laboratorio, muestran que estos podrían corresponder a agregados ordenados, fibrilares, con características amiloides (J.M. Centeno, com. pers.); mas no extenderemos aquí la discusión al respecto.

ANÁLISIS DEL MECANISMO DE UNIÓN E2C-ADN POR MUTAGÉNESIS SITIODIRIGIDA:
EFECTO SOBRE EL MECANISMO CINÉTICO DE INTERACCIÓN

2

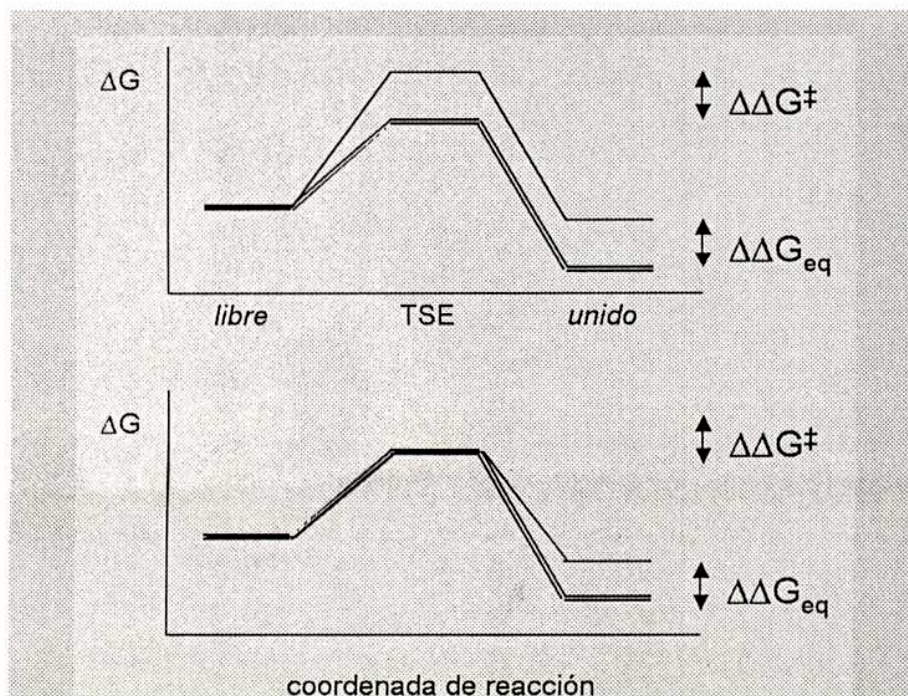
"Never did I think myself as panorama, or as part of a panorama, whether the panorama be human or not. No, never. But molecular biology nowadays invades everything, and if it is metamorphosing into a landscape, what can we do? Paint. This is my intimate, heuristic, molecular, and human conclusion."

André Lwoff

ANÁLISIS DEL MECANISMO DE UNIÓN E2C-ADN POR MUTAGÉNESIS SITIODIRIGIDA: EFECTO SOBRE EL MECANISMO CINÉTICO DE INTERACCIÓN

En última instancia, el camino de reacción de la unión ente macromoléculas queda completamente definido cuando todos los estados conformacionales de las rutas de interacción, incluyendo los elementos libres, unidos, los intermediarios, los estados de transición y todas las constantes de interconversión, están identificados a nivel de cada uno de los aminoácidos o grupos interactuantes. Para esto es necesario recabar información sobre muchos parámetros diferentes involucrados en la estabilidad de la unión y definirlos, en términos estructurales y energéticos, durante todo el proceso de reacción ⁴³. El método de "ingeniería de proteínas" fue desarrollado a principios de la década pasada para el estudio de reacciones de plegamiento proteico ²⁵. Este se basa en la posibilidad de estudiar los intermediarios y los conjuntos de estados de transición (TSE) por medio de la caracterización del efecto causado por mutaciones racionales puntuales sobre las constantes de reacción cinéticas y su relación con la constante de equilibrio de la reacción ²⁵. Como se esquematiza en la figura 5.1, una mutación puntual puede afectar sólo el estado final de la interacción (*unido*), afectando el valor de la energía libre de Gibbs en " $\Delta\Delta G_{eq}$ " (Fig 5.1 abajo). Sin embargo, la mutación puede también afectar el TSE que separa ambos estados (Fig 5.1, arriba), lo cual indica que el residuo sustituido está involucrado en interacciones en el estado de transición. Si la mutación desestabiliza el TSE, la reacción procede más rápidamente ya que la barrera de activación disminuye ($\Delta\Delta G^\ddagger$). De forma análoga, si la mutación estabiliza el TSE, la velocidad de reacción procede más lentamente.

Figura 5.1: Diagrama de energía libre de una asociación entre macromoléculas. En la parte superior se esquematiza el efecto de una mutación que desestabiliza tanto el estado final (*unido*) como el estado de transición (TSE). En la parte inferior, el caso en que la mutación desestabiliza el estado *unido* pero no afecta el estado de transición. La línea gruesa, reacción de la proteína salvaje; línea fina, reacción de una mutante puntual.



EFFECTO DE LAS MUTACIONES PUNTUALES SOBRE LAS ETAPAS DE LA REACCIÓN

Para el caso de la reacción de unión entre E2c y ADN, nos propusimos evaluar el camino cinético de interacción a nivel de cada uno de los aminoácidos de E2c que compone la superficie de interacción con ADN. Para esto utilizamos el conjunto de mutaciones puntuales antes descrito (capítulo 4), y evaluamos el efecto causado por cada modificación sobre cada uno de los eventos cinéticos caracterizados (capítulo 3).

Sobre cada una de las mutantes de E2c se realizaron las siguientes mediciones:

- 1) **Curva de unión en pseudoprimer orden respecto a FI-s35:** con el fin de evaluar los efectos causados por la mutación sobre las reacciones de colisión con ADN, se ensayó la unión a 5nM FI-s35 y cantidades crecientes de cada una de las proteínas, siguiendo el cambio de intensidad de la fluorescencia de FITC en el tiempo, como lo realizado para la proteína salvaje (capítulo 3, Fig 3.1). Se realizaron entre 6 y 8 mediciones para cada proteína. Las k_{on} calculadas se muestran en la Tabla 5.1. (tablas 5.1 y 5.2, seguidas a pag completa)
- 2) **Fluorescencia de Trp:** en el camino cinético de interacción entre E2c y ADN identificamos dos intermediarios que presentan cambios en la fluorescencia de los Trp localizados en el interior del barril β -dimérico (capítulo 3, Fig 3.4). Estos resultados fueron obtenidos utilizando la mutante W339F, que presenta cambios fluorescentes menos complejos que la proteína salvaje. Como todas las mutaciones puntuales realizadas sobre la superficie de interacción se construyeron sobre la proteína salvaje, la comparación la realizamos sobre los eventos fluorescentes identificables de la proteína salvaje. A fin de minimizar las señales fluorescentes provenientes de los eventos de colisión y al mismo tiempo obtener una mayor amplitud en la señal, utilizamos una alta concentración de proteína (100nM), que mezclamos por flujo detenido con 300nM de ADN s35. Todas las curvas fueron ajustadas a decaimientos exponenciales y las constantes cinéticas recuperadas se muestran en la tabla 5.1.
- 3) **Expulsión de solvente de la interfase:** como se propuso en el capítulo 3, la remoción de solvente de la interfase proteína-ADN puede ser monitoreada por la expulsión concomitante del fluoróforo ANS (Fig 3.6). Cada una de las mutantes fue pre-equilibrada en 50 μ M ANS y la velocidad de expulsión medida por el cambio fluorescente del ANS, ante la mezcla de un equivalente de s35. Las constantes de velocidad recuperadas se muestran en la tabla 5.1

4) Cinética de disociación: a fin de evaluar el efecto de las mutaciones sobre los caminos cinéticos de disociación, realizamos ensayos de disociación cinética, como lo descrito para la proteína salvaje (capítulo 3, Fig 3.3). Los resultados se resumen en la tabla 5.2

Tabla 5.1: parámetros cinéticos de cada mutante sitiodirigida para los eventos de colisión y rearreglos conformacionales

	Pseudoprimer orden Fls-35		Fluorecencia de Trp			ANS
	$k_{on} 1 (nM^{-1}s^{-1})$	$k_{on} 2 (nM^{-1}s^{-1})$	$kW1 (s^{-1})$	$kW2 (s^{-1})$	$kW3 (s^{-1})$	$kANS(s^{-1})$
wt	0.87	0.442	27.6 ± 1.5	0.060 ± 0.01	0.0056 ± 0.001	0.0068
A293G	0.2245	$9E-05$	1.9 ± 0.8	0.099 ± 0.03	0.0046 ± 0.003	0.0063
N294D	0.527	0.009	71.4 ± 10.0	0.056 ± 0.01	ND	0.0128
N294A	NM	NM	9.8 ± 5.0	0.091 ± 0.02	0.0114 ± 0.005	0.0042
T295A	0.8533	0.302	16.6 ± 6.8	0.036 ± 0.01	0.0050 ± 0.002	0.0056
K297A	0.6816	ND	16.6 ± 5.0	0.066 ± 0.01	0.0017 ± 0.001	0.0079
K297R	1.292	0.054	31.4 ± 5.0	0.061 ± 0.00	0.0063 ± 0.001	0.0175
C298S	0.794	0.096	25.6 ± 1.5	0.053 ± 0.01	0.0137 ± 0.008	0.0066
C298G	0.45	0.191	ND	0.101 ± 0.05	0.0630 ± 0.002	0.0152
C298A	0.175	0.118	ND	0.069 ± 0.01	0.0054 ± 0.001	0.0154
R300M	ND	ND	ND	0.101 ± 0.02	0.0185 ± 0.005	0.0642
Y301N	0.3597	0.076	69.5 ± 29.4	0.047 ± 0.01	0.0113 ± 0.001	0.006
Y301F	0.7286	0.384	50.6 ± 23.8	0.032 ± 0.01	0.0054 ± 0.002	0.004
Y301A	0.7275	0.194	38.7 ± 10.0	0.070 ± 0.01	ND	ND
R302Q	0.5115	0.07	32.3 ± 6.1	0.039 ± 0.03	0.0044 ± 0.002	0.00257
R302M	0.665	0.211	8.8 ± 5.0	0.222 ± 0.06	0.0043 ± 0.001	0.0065
F303L	0.926	0.401	6.9 ± 1.6	0.027 ± 0.02	0.0017 ± 0.000	0.003
K304A	0.4305	0.037	20.4 ± 12.3	0.049 ± 0.02	0.0038 ± 0.003	0.00686
K305A	0.683	0.277	30.2 ± 5.0	5.900 ± 2.00	0.0640 ± 0.005	0.0026
V324A	0.9072	0.447	47.9 ± 8.2	0.062 ± 0.02	0.0049 ± 0.001	0.0071
K325A	0.4516	0.236	34.5 ± 0.4	0.518 ± 0.03	0.0403 ± 0.009	0.0318
K325R	1.293	0.004	13.5 ± 7.0	0.057 ± 0.01	0.0057 ± 0.010	0.00925
K327A	0.3512	0.272	34.2 ± 1.9	0.045 ± 0.01	0.0062 ± 0.001	0.0096
K327R	0.8331	0.681	34.9 ± 2.8	0.041 ± 0.01	0.0046 ± 0.003	0.0167

Tabla 5.2: parámetros cinéticos de disociación de ADN específico para cada mutante sitiodirigida.

	$k1$ (s ⁻¹)	err $k1$	$k2$ (s ⁻¹)	err $k2$
wt	1.16	0.4	0.32	0.05
A293G	2.30	1.0	0.33	0.05
N294D	33.00	10.0	0.25	0.10
N294A	NM	NM	NM	NM
T295A	3.95	0.1	0.05	0.02
K297A	3.35	0.1	0.41	0.01
K297R	29.80	5.9	5.20	5.37
C298S	4.32	1.2	0.12	0.01
C298G	ND	NA	0.42	0.10
C298A	ND	NA	0.37	0.01
R300M	ND	NA	0.46	0.03
Y301N	6.80	1.5	0.24	0.17
Y301F	1.30	0.3	0.48	0.22
Y301A	21.80	5.0	0.49	0.04
R302Q	6.74	2.6	0.14	0.02
R302M	18.85	0.6	0.69	0.02
F303L	0.65	0.2	0.31	0.04
K304A	28.80	3.4	3.50	0.12
K305A	1.33	0.3	0.72	0.15
V324A	1.49	0.1	0.73	0.02
K325A	ND	NA	0.46	0.10
K325R	0.91	0.1	0.19	0.02
K327A	1.17	0.1	0.74	0.07
K327R	ND	NA	0.27	0.10

EFFECTO DE LAS MUTACIONES SOBRE LA COLISIÓN INICIAL

Como se detalló en el capítulo 3, el mecanismo de unión de la proteína salvaje comienza con una colisión inicial extremadamente rápida, limitada por la difusión, y asistida electrostáticamente. Asimismo, las velocidades no varían sustancialmente cuando la colisión se da con oligonucleótidos de secuencia inespecífica, asignando que el complejo de encuentro formado luego de la colisión tiene una baja (o nula) capacidad de "discriminación de secuencia". A fin de evaluar el efecto de la mutación sobre la colisión inicial, se realizaron curvas de pseudo-primer orden respecto a FI-s35, para cada una de las mutantes puntuales (tabla 5.1).

E2c se une a ADN por dos caminos de unión paralelos (capítulo 3, Fig 3.8). Estas rutas de asociación parten de un equilibrio conformacional de E2c, preexistente a la unión de ADN. Como se planteó hacia el final del capítulo 3, las diferencias en las velocidades de colisión pueden deberse a diferencias en las k_{on} intrínsecas de cada conformero o a la población diferencial de cada una de las poblaciones de conformeros. El efecto de las mutaciones será analizado bajo ambas hipótesis:

1) La mutación no afecta la población de cada conformero

En este caso, las velocidades de colisión dependen solamente del efecto de la mutación sobre la k_{on} de cada conformero. En este marco es posible entonces relacionar las constantes de reacción cinéticas obtenidas al efecto de la mutación sobre la energía del estado de transición ($\Delta\Delta G^\ddagger$) de la colisión inicial, ya que ²⁵:

$$\Delta\Delta G^\ddagger = -RT \ln (k_{mut} / k_{wt})$$

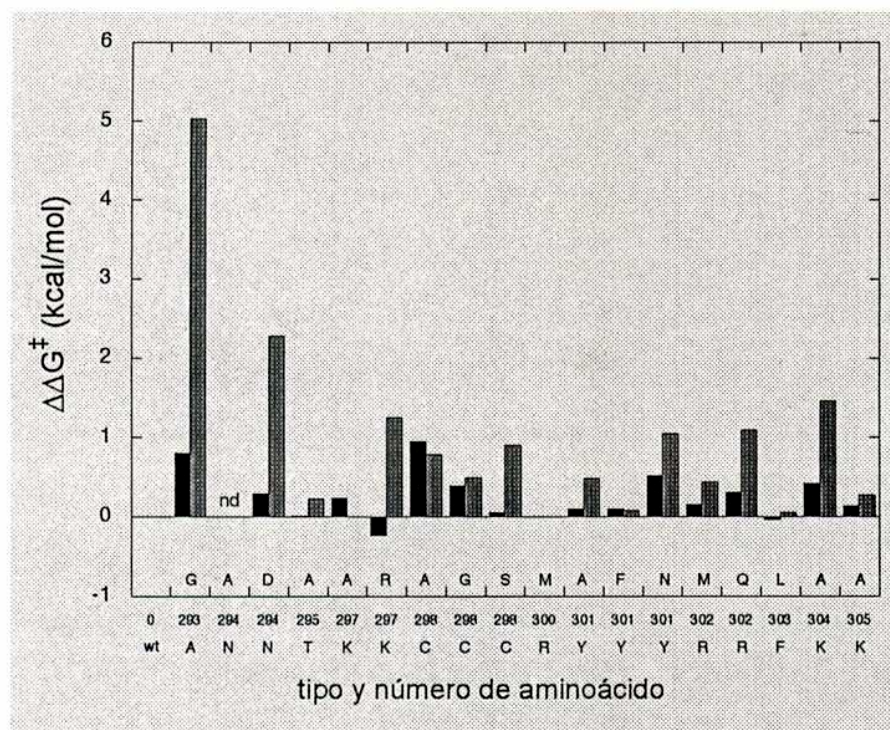
Las diferencias energéticas calculadas se muestran en la figura 5.2. El efecto de la mutación es más drástico sobre la energía de transición de k_{on2} , relacionada a la ruta de asociación con dos intermediarios (capítulo 3, Fig 3.8). Los valores absolutos se encuentran fuertemente afectados, considerando que la barrera energética asignada a este evento es de 4 a 5 kcal/mol para la proteína salvaje (figura 3.12 y tabla 3.3, capítulo 3). Según esta interpretación, la mutación afecta menos a la ruta de asociación de dos estados, siendo la unión del conformero responsable menos perturbado por la mutación. Este resultado está en concordancia con la hipótesis estructural antes planteada, donde la ruta de asociación de dos estados corresponde con un conformero con estructura más *flexible* que, en principio, sería menos sensible a la perturbación por mutaciones puntuales.

Los efectos causados por las mutaciones no son directamente asignables a perturbaciones sobre residuos cargados, como podría esperarse de la delección de cargas superficiales de una colisión difusiva asistida electrostáticamente. Por el contrario, las mutaciones que más afectan a esta etapa son las relacionadas con cambios estructurales detectados por DC (capítulo 4, Fig 4.4).

Figura 5.2: Efecto de las mutaciones sobre la energía libre del TSE de la colisión inicial.

A partir de las constantes cinéticas de colisión, se calculó la diferencia de energía libre de los estados de transición como $\Delta\Delta G^\ddagger = -RT \ln (k_{mut} / k_{wt})$.

Barras negras, $\Delta\Delta G^\ddagger$ calculado para k_{on1} , barras grises, $\Delta\Delta G^\ddagger$ calculado para k_{on2} .



2) La mutación no afecta la k_{on}

Si la mutación introducida no afecta la velocidad de colisión, las diferencias encontradas se deben a la población relativa de las poblaciones de confórmers de E2c. De la misma manera que para la proteína salvaje, la relación entre las velocidades observadas permite estimar la población diferencial de cada uno de los confórmers (tabla 5.3). Cabe mencionar que la asignación de confórmer A o B es arbitraria, siendo por definición el confórmer A el que presenta mayor k_{obs} . Los porcentajes presentados deben ser tomados como aproximados, estimándose un 25% de error acumulado. De todas formas, las mutaciones que más afectan son también las relacionadas con cambios estructurales asignados por DC. Por ejemplo, la mutación A293G y N294D presentan prácticamente un solo confórmer de asociación, indicando que existe una fuerte perturbación en el pre-equilibrio de E2c causado por mutaciones en el N-cap de la hélice de reconocimiento. No hemos encontrado una correlación directa entre los valores de elipticidad y la población relativa de cada confórmer, lo que sugiere que los porcentajes de las poblaciones no son enteramente asignables por DC, o bien porque la asignación de las k_{on} a cada confórmer es arbitraria.

Desde ya, cabe la posibilidad de que tanto la k_{on} como la población relativa de los confómeros se vean afectados por la mutación. En tal caso, no nos es posible asignar los efectos de la mutación sobre los parámetros cinéticos de la reacción de colisión de E2c con ADN. Solo se podría concluir que todas las mutantes medidas presentan más de un evento de colisión bimolecular.

Tabla 5.3: Acumulación relativa de las poblaciones conformacionales de E2c, según la mutación puntual introducida.

	%A	%B
wt	66	34
A293G	99	1
N294D	98	2
N294A	ND	ND
T295A	74	26
K297A	ND	ND
K297R	96	4
C298S	89	11
C298G	70	30
C298A	60	40
R300M	ND	ND
Y301N	83	17
Y301F	65	35
Y301A	79	21
R302Q	88	12
R302M	76	24
F303L	70	30
K304A	92	8
K305A	71	29
V324A	67	33
K325A	66	34
K325R	98	2
K327A	56	44
K327R	55	45

EFFECTO DE LAS MUTACIONES SOBRE LAS FASES DE REARREGLO

Cuando se monitorean los cambios de la fluorescencia de Trp de E2c durante una reacción de asociación con ADN específico a alta concentración, se distinguen 3 fases (k_W 1, k_W 2, k_W 3 en la tabla 5.1). Según lo presentado en el capítulo 3, donde se utilizó la mutante W339F, dos de ellas corresponden a variaciones de los Trp localizados en el centro del barril_ (W317 y W319) y una de ellas al W339, sobre la hélice 2 y expuesto al solvente. Dados los rangos temporales, el evento de expulsión de ANS asociado, y la poca variación que las mutaciones inducen sobre la fase k_W 2, asignamos a ésta como un evento fluorescente local del W339, que ocurre en paralelo a la unión y no es un paso limitante en las rutas de asociación. Es sobre las dos otras fases donde se analizará puntualmente el efecto de cada mutación. La fase k_W 1 la asignamos asociada al primer rearreglo conformacional presentado por el complejo E2c:ADN, analizada en el capítulo 3 (Fig 3.12). La fase k_W 3 corresponde al rearreglo del segundo intermediario que, para la proteína salvaje, ocurre en simultáneo con la expulsión de ANS (Fig 3.8, Fig 3.12, Tabla 5.1). Cabe mencionar que las diferencias entre las velocidades reportadas para la proteína salvaje en el capítulo 3 y en la tabla 5.1 se deben a que, para el análisis mutacional, se agregó 10mM de MgCl2 al tampón de reacción.

La asignación de cada una de las fases cinéticas que presentan las mutantes fue realizada por el orden temporal en el que se suceden y no por las amplitudes relativas de cada una de ellas, dado que el rendimiento cuántico del Trp puede no estar relacionado con las concentraciones relativas de los intermediarios. En los casos en los cuales no se detectaron eventos fluorescentes para una etapa cinética (tabla 5.1, nd), asumimos que la barrera de activación de ese evento es afectada por la mutación. Interpretamos que esto se produce por una desestabilización del TSE, modificándose el camino cinético de forma que esa etapa deja de ser un paso limitante en la unión y por tanto no es detectada.

A partir de las constantes cinéticas observadas para cada uno de los rearreglos conformacionales, calculamos la diferencia en la energía libre de Gibbs ocasionada por la mutación en cada uno de los TSE de las etapas cinéticas. Los resultados obtenidos para los rearreglos detectados por fluorescencia a Trp se muestran en la figura 5.3, los obtenidos para la fase detectada por fluorescencia de ANS, en la figura 5.4.

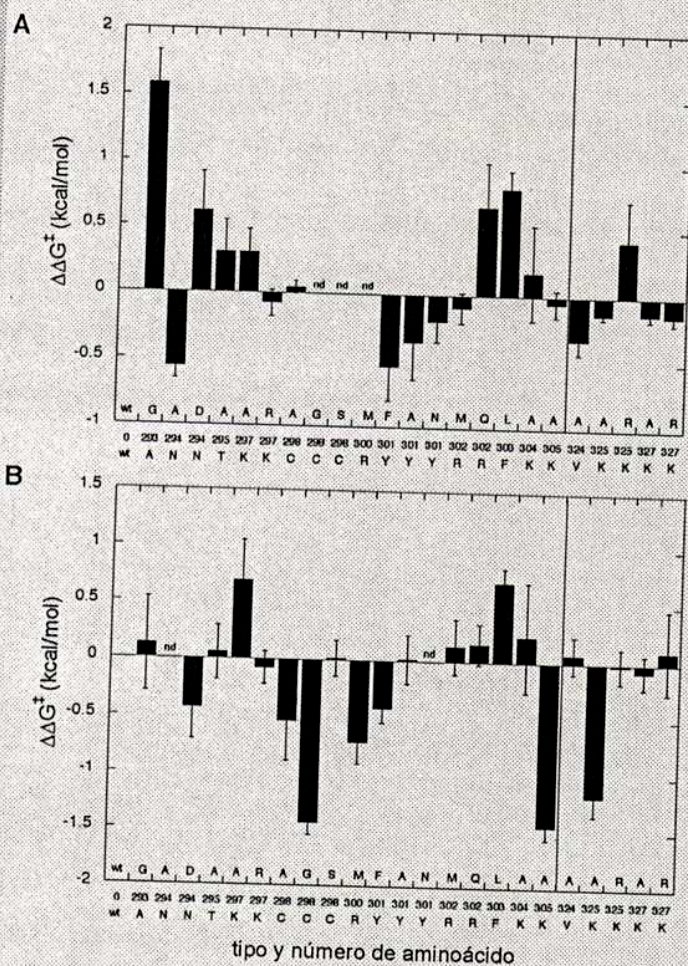
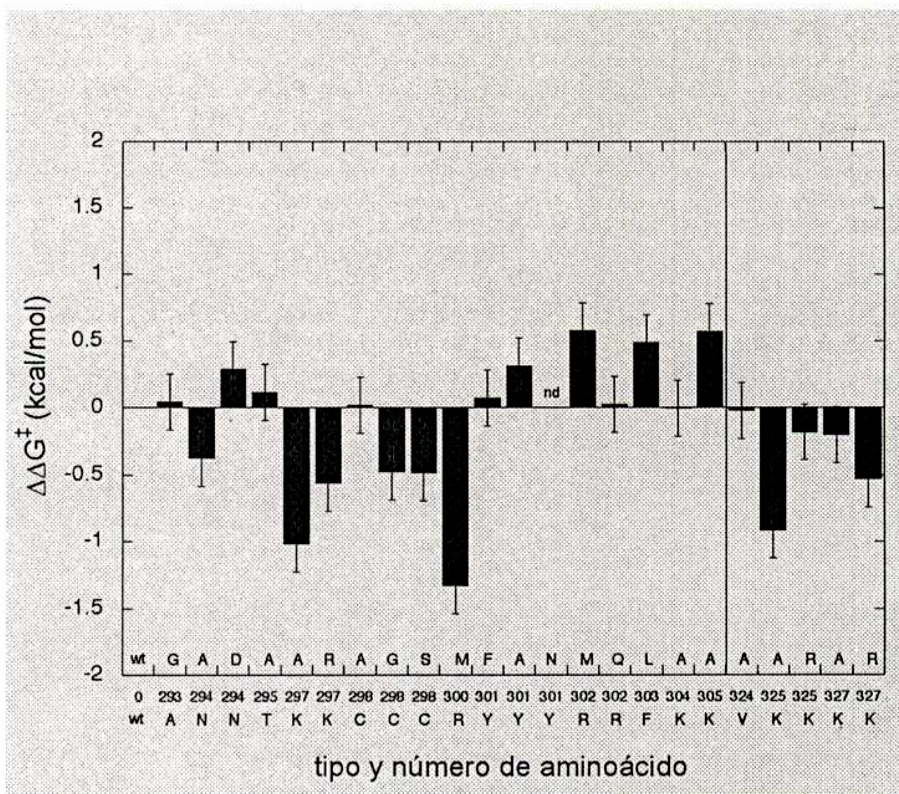


Figura 5.3: Efecto de las mutaciones sobre la energía libre de los TSE de rearrreglos conformacionales de E2c. A partir de las constantes cinéticas k_{W1} y k_{W3} , se calculó la diferencia de energía libre de los estados de transición como $\Delta\Delta G^\ddagger = -RT \ln(k_{mut} / k_W)$. A) Efecto sobre el TSE del primer rearrreglo conformacional (k_{W1}). B) Efecto sobre el TSE del segundo rearrreglo conformacional (k_{W3}). La raya vertical separa las mutaciones realizadas en la hélice1 de las del *loop* $\beta 2$ - $\beta 3$.

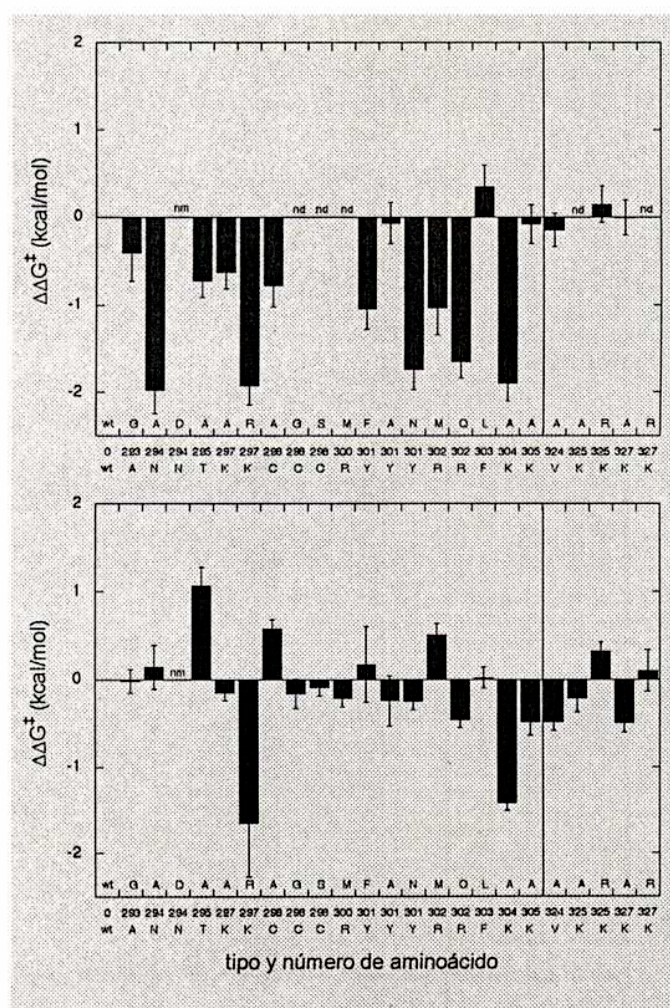
Figura 5.4: Efecto de las mutaciones sobre la energía libre del TSE asociado a la expulsión de solvente de la interfase. A partir de las constantes cinéticas k_{ANS} (tabla 5.1), se calculó la diferencia de energía libre de los estados de transición como $\Delta\Delta G^\ddagger = -RT \ln (k_{mut} / k_{wt})$.



EFFECTO DE LAS MUTACIONES SOBRE LOS EVENTOS DE DISOCIACIÓN

Como se detalló en el capítulo 3, la disociación de E2c de ADN procede por 2 caminos paralelos. Cada uno de ellos puede ser detectado por un cambio de fluorescencia del motivo FITC unido al Fl-s35, cambio que ocurren en paralelo a la disociación de E2c del ADN. A partir de las mediciones presentadas en la tabla 5.2, calculamos la diferencia de energía libre de los TSE de disociación como en los casos anteriores. Los resultados se muestran en la figura 5.5.

Figura 5.5: Efecto de las mutaciones sobre la energía libre de los TSE de disociación. A partir de las constantes cinéticas de disociación (tabla 5.2), se calculó la diferencia de energía libre de los estados de transición como $\Delta\Delta G^\ddagger = -RT \ln (k_{mut}/k_{wt})$. A) Efecto sobre el TSE de la disociación por la ruta de dos intermediarios. B) Efecto sobre el TSE de la disociación por la ruta de dos estados.



ANÁLISIS INDIVIDUAL DEL EFECTO DE CADA MUTACIÓN SOBRE EL CAMINO CINÉTICO

MUTACIONES SOBRE LA "HÉLICE DE RECONOCIMIENTO"

Ala 293: el hecho que la velocidad de reacción de la fase k_{W1} presentada por esta mutante sea más lenta que la observada para la proteína salvaje, indica que la cadena lateral removida participa en la desestabilización del estado de transición que separa el "complejo de encuentro" del intermediario subsiguiente. Al estar ausente la cadena lateral, esta barrera energética aumenta en 1.5 kcal/mol (Fig 5.3A). Dado que este residuo no forma interacciones directas con ADN en el estado "final", la interacción que desestabiliza al TSEw1 es asignable a un requerimiento indirecto sobre el cambio conformacional reportado por este evento. Dado que A293 está localizado en el N-cap de la hélice, y la mutación tiene efectos sobre el plegamiento local de esta hélice (Fig 4.4, capítulo 4), es probable que TSEw1 involucre movimientos locales de plegamiento de esta zona de unión de ADN.

La energía libre del TSEw3 no se ve afectada al remover la cadena lateral de la posición 293 (Fig 5.3B), indicando que esta no es requerida para la transición desde el intermediario 2 al estado final. Para la proteína salvaje, esta fase fue asociada al cambio conformacional coincidente con el evento de expulsión de solvente de la interfase y, como se muestra en la tabla 5.1 y la Fig 5.4, la reacción de expulsión de solvente tampoco se ve afectada por esta mutación. Tomados en conjunto, estos datos sugieren que los cambios conformacionales asociados a la última etapa no involucran grandes cambios conformacionales de la hélice 1, coincidente con lo propuesto en el análisis termodinámico de esta barrera de transición (Fig 3.12, capítulo 3)

En cuanto a los eventos de disociación del ADN específico, la sustitución A293G no presenta cambios respecto a la proteína salvaje, indicando que el metilo de la cadena lateral no forma interacciones en ninguno de los TSE de las rutas de disociación descriptas.

N294: la sustitución de la cadena lateral por un metilo (N294A), aumenta la energía libre del TSEw1, indicando una estabilización de esta barrera de activación (Fig 5.3A). Esto es indicativo que la Asn forma algún contacto que desestabiliza este TSE, un contacto que Ala no puede sustituir. Sin embargo, la mutación por Asp (N294D) restituye esta interacción, e incluso desestabiliza aún más el TSEw1 (Fig 5.3A). Esto sugiere que una de las interacciones que N294 forma en esta instancia es reemplazable por el carboxilato de Asp, y que el grupo amida de Asn forma otra interacción (de alrededor de 0,5 kcal/mol) que desfavorece el pasaje entre el complejo de encuentro y el segundo intermediario.

El efecto sobre el TSEw3 es aún más drástico, ya que la introducción de la carga dada por la sustitución Asn a Asp desestabiliza fuertemente este estado de transición (Fig 5.3B). Este efecto es menor si la sustitución es por Ala, sugiriendo que los rearrreglos conformacionales que se dan durante esta etapa involucran el carácter parcialmente hidrofóbico de la cadena lateral. El mismo efecto se observa en la etapa asociada, la expulsión de solvente, donde se ve que el reemplazo por Asp acelera este evento, que ha de ocurrir por algún rearrreglo no reportado por la fluorescencia de los Trp de E2c.

Las medidas de disociación se realizaron solo para la mutante N294D, que mostró afectar la velocidad de la disociación 1 (k_{dis1} , tabla 5.2, Fig 5.5), asociada al camino de unión de dos intermediarios. Esto está en concordancia con lo presentado por la fase k_W3 , adjudicada a la reacción reversa de la disociación 1, donde la mutación N294D desestabiliza este estado de transición.

T295: la mutación a Ala no afecta las constantes observadas, tanto para los eventos detectados por fluorescencia de Trp (W1 y W3) como de ANS, indicando que los TSE no se ven afectados por la delección de los grupos metilo e hidroxilo de la cadena lateral de Thr (Fig 5.3). Esto permite inferir que los contactos con ADN, atribuidos a este grupo según las estructuras cristalográficas (Fig 4.1C, capítulo 4), se darían en etapas tardías del evento de unión, posterior al último evento conformacional detectado. En línea con esto último, la mutación a Ala afecta los eventos de disociación, estabilizando el TSE de la ruta de dos estados (Fig 5.5 B) y desestabilizando el de la otra ruta (Fig 5.5A).

K297: la sustitución por Ala estabiliza levemente TSEw1, y el reemplazo por otro grupo con la misma carga (Arg), restituye el efecto a niveles salvajes. Esto sugiere que es la carga del motivo amino de la Lys297 el responsable de la estabilidad de TSEw1, y que en este evento no hay una restricción conformacional importante, ya que el reemplazo por un grupo voluminoso resulta en niveles de activación semejantes al salvaje.

El mismo efecto descrito para TSEw1 se aplica al análisis de TSEw3. Es probable entonces que, durante ambos estados de transición, K297 establezca una interacción iónica, aunque no necesariamente las mismas que las reportadas cristalográficamente.

Ambos reemplazos tienen efectos desestabilizadores sobre los TSE de los eventos reportados por la expulsión de ANS de la interfase de unión. Dado que, a diferencia de la proteína salvaje, estos eventos tienen lugar en una escala temporal menor que la fase k_W3 , el evento de expulsión de solvente ha de ocurrir por una vía diferente que la de la proteína salvaje.

Acorde con el efecto que estas mutaciones tienen sobre la reacción de unión al equilibrio (Fig 4.5, capítulo 4), los eventos de disociación se encuentran fuertemente acelerados, indicando TSEs' desestabilizados para ambos caminos de disociación. Este efecto resulta más pronunciado en la sustitución por Arg sugiriendo que, para estas etapas y a diferencia de las etapas de asociación, existen importantes restricciones conformacionales alrededor del residuo 297, inserto en los surcos mayores del ADN.

C298: la interacción que la cadena lateral de Cys realiza en TSEw1 es reemplazable solo por Ser (Fig 5.3A). Las otras sustituciones ensayadas (Ala y Gly) afectan fuertemente este estado de transición, no detectándose cambios fluorescentes asociados. Esto nos permite inferir que, a diferencia de lo observado en el equilibrio (Fig 4.5, capítulo 4) es la capacidad de la Cys de participar en uniones de hidrógeno transiente (reemplazable por Ser) la que estaría involucrada en este rearrreglo conformacional.

La energía libre de transición de la fase k_W3 se afecta cuando la mutación en la posición 298 es a Gly y más levemente a Ser (Fig 5.3B). Por el contrario, ésta no se encuentra afectada cuando la mutación es a Ala. Esto apunta sobre el carácter hidrofóbico de la Cys como característico de las Interacciones que realiza en TSEw3.

La expulsión de solvente de la interfase no se encuentra afectada cuando la mutación sobre C298 es a Ser (Fig 5.4). De la manera análoga a lo descrito para el efecto sobre TSEw1, las interacciones en las que Cys participa durante esta etapa son sólo reemplazables por Ser.

Cabe destacar aquí que, a pesar de ocurrir en la misma escala temporal para la proteína salvaje, los eventos reportados por las fases k_W3 y k_{ANS} son distintos, en tanto las interacciones que realiza la Cys de la hélice de unión son marcadamente diferentes en ambas etapas.

R300: El reemplazo de este grupo por Met desestabiliza el TSEw1, haciendo que deje de ser un paso limitante para esta etapa cinética (Fig 5.3A). Proponemos entonces que Arg300 forma una interacción en TSEw1 que no es reemplazable por Met, probablemente una unión salina con los grupos fosfato del ADN.

Por otro lado, el reemplazo por Met desestabiliza tanto del TSEw3 como al TSE reportado por k_{ANS} . Indicando que este grupo realiza Interacciones que estabilizan ambos estados de transición (Fig 5.3B, Fig 5.4). Es posible que el efecto de la mutación no sea directo sobre ambas etapas, sino que el hecho de afectar a una de ellas (k_W3) conlleve una desestabilización de la etapa concomitante (k_{ANS}).

La delección de la carga sobre Arg300 afecta fuertemente la ruta de disociación de dos intermediarios, en concordancia con la reversa, TSEw3. Por otro lado, no afecta significativamente la disociación por la ruta de dos estados.

Y301: sólo la sustitución de la Tyr por Asn afecta sustancialmente el TSE del primer rearrreglo conformacional (Fig 5.3A). Esto es indicativo de que ni el OH de Tyr, ni el anillo aromático están involucrados en las interacciones que este residuo realiza en esta etapa cinética. La estabilización del TSEw1 causada por Asn puede entonces deberse a la inclusión del grupo amida, que presumiblemente modifique a otras interacciones polares que pudieran ocurrir.

Durante el estado de transición del segundo rearreglo conformacional, el efecto producido por la mutación a Ala, en comparación con la sustitución por Phe, permite atribuir al anillo aromático de Y301 un rol en la estabilización de este TSE. Este efecto también se refleja en la fase asociada de expulsión de solvente, caracterizada por k_{ANS} (Fig 5.3B, Fig 5.4).

El estado de transición de la ruta de disociación de dos estados no se afecta por la mutación del residuo Tyr. Sin embargo, y en relación a lo descripto para TSEw3 y para el efecto al equilibrio, existe un marcado efecto desestabilizador cuando sólo cuando se remueve el anillo aromático de la Tyr.

R302: las interacciones que la Arg de esta posición contribuye al TSE del primer rearreglo conformacional son reemplazables por Gln, indicando que la carga de la Arg no está involucrada en la estabilización del TSEw1. Sin embargo, la sustitución de Arg por un grupo hidrofóbico voluminoso (Met) desestabiliza a este estado de transición (Fig 5.3A). Tomados en conjunto, estos resultados sugieren que es el carácter parcialmente hidrofóbico del residuo Arg el involucrado en este rearreglo conformacional.

Ninguna de las sustituciones realizadas sobre Arg 302 afectan el evento reportado por la fase k_W 3, indicando que la cadena lateral de la Arg no estaría involucrada en interacciones en el TSE de esta etapa (Fig 5.3B). Durante el evento de expulsión de solvente, el efecto de Arg es reemplazable sólo por Met, siendo posible adjudicarle interacciones de tipo hidrofóbico este residuo en esta etapa (Fig 5.4). La inclusión de un grupo polar (Gln) estabiliza el TSE de este evento.

En el caso del evento de disociación por la ruta de dos intermediarios, ambas mutaciones afectan el TSE, presumiblemente por afectar el estado basal del complejo al equilibrio. En cambio, la mutación por Gln estabiliza el TSE de la ruta de disociación por dos estados, sugiriendo que este residuo realiza interacciones en este estado, presumiblemente mediadas por sus características polares.

F303: la mutación de este residuo a Leu es la única que mapea en la cara interna de la hélice de unión. El hecho de que ejerza efectos estabilizadores sobre TSEw1 y TSEw3, es indicativo que durante estas etapas existen rearreglos de la hélice de unión, no involucrados directamente con el contacto de ADN (Fig 5.3). Las fluctuaciones conformacionales que el reemplazo por Leu afecta, serían requeridas para bajar las barreras energéticas que separan los intermediarios de asociación. En línea con esto, también se ve afectada levemente la fase de expulsión de solvente, k_{ANS} (Fig 5.4).

Contrariamente a los efectos observados para los eventos de asociación, la mutación a Leu en la cara interna de la hélice1, no afecta las rutas de disociación.

K304: la mutación a Ala de este residuo, localizado hacia el N-terminal de la hélice de unión, no afecta ninguno de los rearreglos conformacionales de asociación (Fig 5.3, Fig 5.4). Esto es indicativo de que, en los estados de transición reportados, la Lys no participa de interacciones que no puedan ser reemplazadas por Ala.

El efecto que la mutación K304A tiene sobre la unión al equilibrio se refleja en el efecto desestabilizador de los estados de transición de las disociaciones. En ellas se distingue un marcado efecto sobre ambas rutas, sugiriendo que las interacciones en las que K304 participa se forman luego de los TSEs de asociación descriptos.

K305: la mutación a Ala del residuo más extremo de la hélice de unión no afecta el TSE del primer rearreglo conformacional (Fig 5.3A). En cambio, se distingue que la Lys 305 juega un rol desestabilizador del TSE del segundo rearreglo conformacional (TSEw3), bajando la barrera de activación y permitiendo que la reacción proceda al estado "final" (Fig 5.3B).

Esta mutante permite distinguir temporalmente el evento de rearreglo k_{W3} y la expulsión de solvente, ya que para tener efectos inversos en ambos TSE (Fig 5.4). En este caso, la mutación resulta desestabilizar el TSE de la expulsión de solvente de la interfase, desacoplandola con el efecto estabilizador sobre TSEw3. Esto es indicativo de que el residuo K305 realiza alguna interacción, no reemplazable por Ala, que enlentece la expulsión de solvente de la interface.

Los efectos de la mutación K305A sobre los eventos de disociación son bajos en el caso de la disociación por la ruta de dos estados y nulos en la de dos intermediarios, indicando que estos estados de transición tienen características estructurales diferentes (Fig 5.5)

MUTACIONES SOBRE EL LOOP $\beta 2$ - $\beta 3$

V324: la mutación de este residuo por Ala causa un efecto sutil, pero medible, estabilizando el TSEw1, y no afecta el TSEw3 (Fig 5.3). Esto representa una evidencia directa que residuos no cargados del loop $\beta 2$ - $\beta 3$ participan el rearreglos conformacionales que se dan durante la unión de ADN, y no en etapas finales de los caminos de unión, como la reportada por k_{W3} o por k_{ANS} (Fig 5.4). Por otro lado, la mutación a Ala afecta sólo y mínimamente la ruta de disociación de dos estados (Fig 5.5).

K325: ninguna de las mutaciones realizadas sobre este residuo del *loop* afectan la fase k_{W1} (Fig 5.3A). Esto nos permite inferir que la cadena lateral de Lys325 no esta involucrada en interacciones del TSEw1. Sin embargo, la delección de la carga por la mutación a Ala desestabiliza el TSEw3 y a TSEans (Fig 5.3B, Fig 5.4). Dado que el efecto no se observa en la mutante a Arg, es plausible suponer que la interacción responsable es de carácter iónico. En concordancia con lo observado al equilibrio, estos resultados nos permiten proponer que la zona del *loop* no presenta restricciones conformacionales que requieran una geometría estricta para este tipo de uniones.

Por otro lado, únicamente la falta de la carga en la posición 325 afecta la disociación por la ruta de dos Intermediarios. La ruta de dos estados no se encuentra afectada (Fig 5.5).

K327: a pesar de compartir características químicas y cercanía estructural, los efectos de las mutaciones sobre los eventos cinéticos en K327 difieren de K325. Ninguna de las mutaciones realizadas sobre K327 afectan significativamente las energías libres de los TSE de los rearreglos en E2c reportados por fluorescencia de Trp (Fig 5.3). Esto es indicativo que la cadena lateral de K327 no forma interacciones en estos estados de transición.

Se distingue un efecto desestabilizador leve sobre el TSE del evento de expulsión de solvente, sólo cuando la sustitución es a Arg (Fig 5.4). De forma análoga, la disociación por la ruta de dos Intermediarios se ve afectada por esta sustitución, y no causa efecto sobre la disociación de dos estados (Fig 5.5).

Los eventos descritos aquí para cada mutante serán integrados con los de los capítulos precedentes en la discusión general (capítulo 6).

DISCUSIÓN GENERAL Y CONCLUSIONES

La oscuridad no me preocupa. Me preocupa la luz. La oscuridad es solamente ausencia de luz. Pero la ausencia sí me preocupa. La preocupación no. Me es indiferente. Sin embargo, la indiferencia me preocupa muchísimo. La considero una actitud vergonzosa. Aunque la vergüenza no me preocupa. Ante sí, me preocupaba. Pero a mí me da lo mismo el antes y el después; mi vida no es un desarrollo tendiente a la nada. Por eso la nada no me quita el sueño. El sueño, en cambio, es algo que sí me interesa. A veces me quedo toda la noche despierto, pensando en eso. No llego a ninguna conclusión, pero las conclusiones me exasperan. Prefiero los puntos de partida. No por la partida; por los puntos. Siempre trato de acumular puntos. No por lo puntos en sí; es por la acumulación. La acumulación entendida como una cosa sola, no como un cúmulo de cosas. Los cúmulos, yo, si pudiera, los disgregaría. Las cosas tienen que ir separadas; no juntas. Juntas forman otras cosas, y eso trae complicaciones. Aunque yo a las complicaciones no les tengo miedo. Lo que me asusta es lo simple. Lo simple no se sabe de dónde sale; ahí es donde está el misterio. Aunque los misterios, por suerte, no me interesan. Me interesa la suerte. Qué desgracia. Porque la suerte es siempre escasa. Y si dijera que no me preocupa la escasez, mentiría. Pero mentir no me preocupa. A mí me preocupa la verdad. Cuando miento no tengo problema; puedo decir cualquier cosa. Aunque sea verdad, no importa, porque la digo de mentira. Pero cuando hablo con la verdad, tengo que andar con más cuidado. Por las dudas, en esos casos digo lo menos posible. Y después me desdigo, así cubro dos posibilidades. Pero no es que me quiera cubrir. Yo hago todo a la intemperie. Y si no hay luna, mejor. A mí me gusta la oscuridad. (Da capo).

Leo Masliah

En el presente trabajo de tesis se estudiaron en detalle varios aspectos del mecanismo de unión y reconocimiento entre una proteína y su sitio específico de unión al ADN. El sistema elegido fue el de la interacción del dominio C-terminal de la proteína E2 del *papillomavirus* humano cepa 16 (hE2c), con el "elemento de respuesta E2" (s35). La altísima complejidad de las interacciones entre proteínas y ADN *in vivo* depende de parámetros bioquímicos tales como pH, concentración de sales, estados redox, actividad de agua, "crowding" molecular y, más relevante aún, depende de las interacciones con otras proteínas, las modificaciones postraduccionales, etc. Nos propusimos analizar la interacción en condiciones en las cuales las variables del entorno fueran mantenidas al mínimo, de forma que fuera posible extraer datos fisicoquímicos precisos de la interacción. Confiamos en que la interpretación de éstos en términos termodinámicos provee un marco para relacionar las estructuras proteicas con su funcionalidad biológica. Lograr comprender los arreglos estructurales y energéticos utilizados en la naturaleza podría conducir al desarrollo de métodos racionales que permitan modificar o diseñar proteínas *de novo* que controlen el reconocimiento específico y la expresión génica.

MECANISMO DE UNIÓN E2c-ADN AL EQUILIBRIO

Contando con la posibilidad de obtener los elementos de estudio con alto grado de pureza y con un ensayo de unión en solución, reversible y no invasivo, caracterizamos el mecanismo de interacción al equilibrio entre el dominio C-terminal de E2 de HPV16 (hE2c) y ADN. A pH 7.0 y 0.2 M NaCl, la estequiometría de reacción encontrada fue de dos dímeros de E2c por sitio de E2 canónico, contenido en un oligonucleótido de 18pb (s35). Esto fue evidenciado por métodos espectroscópicos como el cambio anisotrópico de la fluorescencia de un motivo FITC unido al oligonucleótido (Fig 2.1), fluorescencia intrínseca de los Trp de hE2c (Fig 2.2), desplazamiento del fluoróforo bis-ANS (Fig 2.3), y diroísmo circular (Fig 2.7). Por otro lado, en ensayos de retardo de movilidad electroforética se observó una segunda especie, cuya aparición es concomitante a la desaparición de la primera banda de retardo, atribuida al complejo 1:1 hE2c-ADN (Fig 2.4). Se intentaron realizar separaciones de los complejos 1:1 y 2:1 por columnas de filtración molecular pero, en las condiciones de unión utilizadas, la proteína queda retenida en la matriz teniendo que aumentar la fuerza iónica para eluiría, variable que influye fuertemente la unión E2c-ADN (Fig 2.9).

Disminuyendo la concentración de los reactantes en el ensayo de unión, pudimos separar los dos eventos y calcular las constantes de disociación al equilibrio (K_D). Para la unión 1:1 la K_D medida es de 0.18 ± 0.2 nM, en el rango de valores generalmente reportados para interacciones específicas proteína-ADN, y estimados como fisiológicamente relevantes ^{47, 57}. Para el segundo evento de unión obtuvimos una K_D de 1.8 ± 0.4 nM, demasiado afín para ser considerada como "no-específica" dado que para oligonucleótidos con secuencias no rela-

cionadss (EBNA e iset) medimos constantes tres órdenes de magnitud más débiles (tabla 2.1). No hay evidencias de que E2c forme tetrámeros en solución, y la disociación del dímero en estas condiciones esta en el orden subpicomolar ^{72, 73}, por tanto proponemos que los efectos observados se deben a dos *modos* diferentes de unión a ADN presentados por el dímero E2c de HPV16 (Fig 2.10).

Por otro lado, la unión a ADN del homólogo E2c de BPV-1 (bE2c) presenta una estequiometría de 1:1 con todos los oligonucleótidos ensayados (Fig 2.6). La unión de sitios E2 específicos es 7 veces más debil que la de hE2c, aún cuando se la ensaya con oligonucleótidos cuya secuencia se deriva de sitios de BPV-1 (cognatos de BPV1). La discriminación de secuencias cognatas es 50 veces mayor para la proteína de HPV16, que a su vez presenta un índice de discriminación por sitios inespecíficos 180 veces mayor que bE2c (tabla 2.1). Entonces, a pesar su alta similitud de secuencia, particularmente en las zonas de unión a ADN (90%), y de presentar una topología general de plegamiento idéntica (Fig 1.5, Fig 4.3), estas proteínas presentan diferencias sustanciales en su capacidad de unión y reconocimiento de ADN.

En general, se cree que la baja afinidad de una interacción específica podría estar compensada con una mayor cantidad de sitios de unión en una región particular del genoma. A su vez, se ha visto que cuanto mayor es el componente electrostático de la Interacción, ésta resulta ser más débil y menos específica ^{47, 48}. Hay 17 sitios E2 en el genoma de BPV-1 y 4 sitios E2 en el de HPV16 y, en concordancia con esto, el complejo formado por bE2c:ADN es más sensible a la fuerza iónica que el de hE2c (Fig 2.9). Una comparación detallada de las estructuras cristalográficas resueltas para ambos dominios muestra que hE2c tiene una menor densidad de carga (positiva) en la superficie de interacción con ADN que la contraparte de BPV1. Hines *et al.* propusieron que esto podría estar relacionado con la capacidad diferencial de ambas proteínas de curvar el sitio E2 ³⁹, lo que a su vez está en relación con las diferencias en la estructura cuaternarias que ambos dímeros presentan (Fig 1.5). En este sentido, se ha reportado que la sola neutralización asimétrica de cargas de los grupos fosfato puede inducir una curvatura del eje central de la doble hélice ⁴⁹. Por otro lado, la secuencia de bases influye la topología y la flexibilidad de un B-ADN, siendo las secuencias "TpA" las que permiten mayores grados de curvatura ⁴⁰. Teniendo en cuenta que la capacidad de discriminación de secuencia de hE2c depende de la flexibilidad del sitio, esto puede relacionarse con el enriquecimiento en pares T:A encontrado en las secuencias separadoras de los hemisitios de HPV16; que son las que aparecen curvadas en las estructuras cristalográficas resueltas (Fig 1.4 y Fig 4.3). Así, la estructura del ADN cuando se encuentra unido a E2c ha de ser diferente para los complejos de hE2c y bE2c, algo que se corresponde con los cambios de dicroísmo circular de ADN presentados por ambos complejos (Fig 2.8).

Tomados en conjunto, estos datos sugieren que las diferencias encontradas entre los homólogos de BPV1 y HPV16 se deben a un mecanismo adicional de *lectura indirecta* de la secuencia de bases del sitio E2, particularmente influido por la flexibilidad del separador de los hemisitos. Según este modelo, estas proteínas presentan una capacidad de discriminación sitio-específica que depende de las contribuciones electrostáticas de la unión. Cabe preguntarse entonces si esto puede estar relacionado con uniones electrostáticas de la proteína sobre el separador de los hemisitos. Según la estructura cristalográfica del complejo E2c-ADN de BPV1, el *loop* que conecta las hebras $\beta 2$ - $\beta 3$ forma 4 contactos salinos con los grupos fosfato del separador de los hemisitos (Fig 1.4B, Fig 4.3). Por el contrario, el co-cristal de E2c de HPV18-ADN no presenta densidad electrónica en esta región (Fig 1.4A, Fig 4.3). Si nos conformamos con estas visiones estáticas, esto podría sugerir que serían los contactos del *loop* los que influyen en la flexibilidad del separador y por tanto la capacidad de discriminación de secuencia por *lectura indirecta*, ya que los contactos aminoácido-base que constituyen la *lectura directa* están altamente conservados entre ambas cepas.

Realizamos 5 mutaciones puntuales sitio-dirigidas sobre 3 aminoácidos del *loop* $\beta 2$ - $\beta 3$ de hE2c. Todas las mutaciones retienen la capacidad de unir ADN en forma específica. Las mutaciones de los residuos cargados (K325, K327) a Ala afectan la energía libre de unión en alrededor de 0.5 kcal/mol por enlace. Sin embargo, el reemplazo de Lys por aminoácidos que conservan la carga positiva (Arg) restituye la afinidad de la interacción a los niveles de la proteína salvaje, sugiriendo que estos aminoácidos interactúan sólo electrostáticamente con ADN, muy probablemente contactando la zona del "separador" de los hemisitos (tabla 4.1). En la actualidad no contamos con una estructura resuelta para el complejo con ADN de E2c (HPV16) que nos permita contrastar estos resultados. De todas formas, considerando que pueden existir varias etapas de unión entre la "proteína libre" y el "complejo proteína-ADN", no es estrictamente requerido que se observen estos contactos en los estados "finales" de la unión, ya que pueden ser transientes e influir sobre las etapas intermedias, afectando la unión al equilibrio sin que se detecten diferencias entre los estados "inicial" y "final" que son los que se obtienen de las resoluciones cristalográficas.

DISECCIÓN DE LAS CONTRIBUCIONES A LA ENERGÍA LIBRE DE UNIÓN

Con el fin de evaluar las contribuciones energéticas de la interacción E2c-ADN, realizamos mutaciones sitiodirigidas sobre todos y cada uno de los aminoácidos de la hélice de reconocimiento de E2c. Para la asignación racional de las energías involucradas se debe probar si las sustituciones afectan el estado de plegamiento de la proteína y la estabilidad global, ya que esto alteraría el nivel energético inicial de la interacción.

Hay dos tipos de mutaciones sobre la hélice 1 que afectan el estado de plegamiento de E2c. Unas son las que se localizan sobre la cara interna de la hélice, contactando el barril_ (L296, L299). Las mutaciones de éstos residuos por Ala impidieron la expresión recombinante consistente de la proteína, presumiblemente por desestabilizar el "core" hidrófobo presente en esa zona. Las otras mutaciones que afectan el plegamiento son las que modifican el N-cap de la hélice 1 (A293, N294). El reemplazo de A293 por Gly desestabiliza la hélice de reconocimiento, un efecto frecuentemente observado sobre las estructuras helicoidales por el reemplazo de residuos a Gly ²³ (Fig 4.4). La desestabilización causada hace que el plegamiento local converja en otro mínimo estructural, uno que también es posible observar en la proteína salvaje si se realiza un tratamiento térmico suave (Fig 4.4B). Esto está en estrecha concordancia con la existencia de dos poblaciones de *conformeros* de E2c, previos a la colisión con ADN, presentados en el mecanismo cinético de interacción (capítulo 3, Fig 3.8), y sugiere que tanto la mutación A293G como el tratamiento térmico suave, desplazan este equilibrio conformacional (capítulo 5, tabla 5.3).

Resultados de nuestro laboratorio realizados por dinámica de relajación magnética de los protones amídicos de E2c muestran un alto grado de movilidad en las posiciones 293 y 294 respecto a los otros residuos de la misma hélice (Nadra *et. al.*, enviado). Esto es una medida directa de las fluctuaciones conformacionales que presenta la hélice 1 de E2c que, junto con los efectos presentados por la mutación de las cadenas laterales de éstas posiciones, indican que el N-cap de la hélice podría corresponder a un "bisagra" conformacional en E2c operando sobre la estabilidad de la hélice de reconocimiento. Es de esperar que los cambios conformacionales que involucren a la hélice 1 modifiquen fuertemente la capacidad de unión/reconocimiento de sitios E2, por lo que afectaría las funciones biológicas mediadas por E2c (ver más abajo).

Para las sustituciones que no afectan el estado de plegamiento pudimos medir la contribución energética de cada uno de los aminoácidos. Los valores absolutos obtenidos están en concordancia con los medidos en otros sistemas proteína-ADN, donde se estima que una unión de hidrógeno contribuye entre 0,5 y 1 kcal/mol ⁴³. Cabe destacar que los valores energéticos de estas interacciones no-covalentes varían órdenes de magnitud según los pares interactuantes y la geometría del enlace. Por cálculos de mecánica cuántica se ha estimado que los valores típicos pueden oscilar entre 0.1 y 10 kcal/mol ⁴⁴. Para las interacciones salinas con los grupos fosfato obtuvimos valores entre 0,3 y 1 kcal/mol, en concordancia con lo asignado para este tipo de interacciones para el sistema EcorRI-ADN ²⁴. Si consideramos que estas uniones

son las únicas presentes en los complejos de E2c con oligonucleótidos inespecíficos, la sumatoria de las energías de los posibles enlaces salinos de E2c y ADN es de 7.5 kcal/mol, en excelente concordancia con el valor medido para la interacción de E2c con un oligonucleótido de secuencia inespecífica (~ 7 kcal/mol, tabla 2.1).

Todos los valores energéticos asignados están cercanos al valor de la energía térmica a 25 °C (0,6 kcal/mol). Esto sugiere que, a temperatura ambiente, las uniones de la interfase son lábiles, por lo tanto existiría un alto componente dinámico asociado a la interfase de unión. En concordancia con esto, experimentos preliminares de intercambio protón/deuterio del complejo E2c-ADN, anticipan que los protones de la interfase intercambian rápidamente con el solvente (Ale Nadra, com.pers.).

En varios estudios de interfases macromoleculares realizados utilizando estrategias de mutagénesis sitiodirigida, se identificó que aminoácidos individuales pueden contribuir a la energía libre de asociación desproporcionadamente al resto de las sustituciones ¹⁰⁹. A estos residuos se los denomina "hot-spots" y los contactos que establecen son caracterizados como *claves* de las interfases ¹⁸. Dado que el estudio mutagénico que realizamos sobre E2c cubre más del 90% de la superficie de interacción, y que ninguna delección de cadena lateral contribuye desproporcionadamente a la Interacción, nos es posible afirmar que no hay "hot-spots" en la interfase de unión E2c-ADN. Estudios estadísticos de la distribución de "hot-spots" en interfases muestran que estos están preferencialmente distribuidos en regiones deshidratadas ⁸ y, como muestran las estructuras cristalográficas de los complejos, no hay regiones de estas características en la interfase E2c-ADN. De forma análoga, la presencia de "hot-spots" puede ser interpretada como un efecto producido por un alto grado de cooperatividad en la Interacción, en el sentido que la disrupción de un contacto puntual afecta sustancialmente la red total de interacciones. Dicho de otra forma, si hay "hot-spots", las contribuciones energéticas de los residuos no son independientes entre sí. En concordancia con la ausencia de "hot-spots" en la interfase E2c-ADN, mostramos que los efectos causados por las sustituciones individuales son aditivos, indicando que las contribuciones energéticas de los residuos individuales son independientes y por lo tanto no habría una cooperatividad significativa en la unión cuando se la analiza al equilibrio (Fig 4.5, tabla 4.2). Un efecto aditivo de las diferencias energéticas de sustituciones puntuales fue observado en la Interacción del represor *lambda* y ADN específico ^{71, 79}. Hasta donde sabemos, ésta es la única otra interacción proteína-ADN analizada energéticamente en solución, por lo que es tentador especular con que el efecto aditivo sea una característica general del reconocimiento sitio-específico.

Por lo antedicho, nos es posible confiar en que las mediciones realizadas con las mutantes sitiodirigidas reflejan los valores de energía libre de la delección de un grupo funcional en particular. Esto equivale a decir que los efectos de las mutaciones no se propagan por la interfase de asociación, en tanto que la energía de unión está delocalizada por el sistema de interacción en conjunto. En esta instancia no nos es posible asignar cuál es el componente energético (entálpico

o entrópico) que cada mutación afecta. Es de esperar que cada residuo contribuya en una baja proporción a cada factor, y al mismo tiempo es probable que se de una compensación entrópico-entálpica como las comúnmente observadas para reacciones que involucran macromoléculas ^{22, 23}.

Para algunos aminoácidos clave de la interfase realizamos más de un reemplazo de cadena lateral, con vistas a asignar componentes energéticos a nivel de grupos funcionales. Las diferencias energéticas observadas para este tipo de reemplazos están en estrecha concordancia con los datos estructurales de la interfase E2c-ADN, validando indirectamente el abordaje metodológico utilizado (Fig 4.8). De esta manera, pudimos observar que la misma sustitución aminoacídica afecta diferencialmente la energía libre de unión según la posición estructural en la que se encuentre. Por ejemplo, el reemplazo de Lys por Ala o Arg ejerce efectos muy diferentes según se realicen en la posición 297 o 325 de E2c (Fig 4.5). Interpretamos que esto se debe a que diferentes regiones de la interfase presentan distintas restricciones conformacionales. Así, es posible que un reemplazo aminoacídico que conserva una característica química particular no afecte la energía libre de unión si se lo realiza en regiones de la proteína que presentan menores restricciones conformacionales, i.e. el *loop* $\beta 2$ - $\beta 3$.

Si se representa el efecto causado por las mutaciones sobre la energía libre de unión en la estructura del complejo E2c-ADN se distinguen 2 regiones diferentes (Fig 4.6). Los residuos que se exponen hacia el centro del sitio de unión presentan una mayor contribución a la energía libre total de interacción que los ubicados sobre el lado externo del mismo. No hemos encontrado una correlación entre las contribuciones energéticas y el número de contactos que cada residuo realiza, lo cual no es extraño pues ya hemos visto que los valores de las energías no-covalentes varían considerablemente entre sí. A diferencia de lo reportado en estudios plegamiento y de interfases proteína-proteína ^{12, 24}, no encontramos que exista una correlación entre las energías libres medidas y la superficie de van der Waals de cada aminoácido, ni con el volumen comprometido, ni con la diferencia de superficie accesible al solvente que cada residuo oculta en la interacción. Es posible entonces concluir que el efecto posicional descrito no es reducible a las contribuciones superficiales de cada residuo particular, sino que sería un efecto causado por el sistema conjunto de la red de interacción. Esto podría estar relacionado con grados de movilidad diferencial en cada zona, lo cual contribuiría a los componentes entrópicos de la interacción, y por tanto a la energía libre total. Estudios de calorimetría de titulación isotérmica de la unión de ADN de cada una de las mutantes puntuales permitirán diferenciar las contribuciones energéticas y poner a prueba esta hipótesis.

COLISIÓN INICIAL E2c- ADN Y FORMACIÓN DEL "COMPLEJO DE ENCUENTRO"

La caracterización del mecanismo cinético de interacción nos permitió distinguir las etapas por las cuales transcurre la asociación entre E2c y ADN, y asignar las contribuciones que cada residuo de E2c realiza en las diferentes fases de asociación.

La colisión inicial es extremadamente rápida, limitada sólo por la difusión de las macromoléculas en solución. Dadas las características electrostáticas del par interactuante (proteína básica, ADN ácido), es de suponer que la velocidad límite de difusión se encuentre aumentada por una atracción electrostática⁹⁹. Ha sido reportado que los campos electrostáticos generados por la disposición de cargas superficiales pueden actuar a gran distancia (50-70 Å) y acelerar los eventos difusivos, pudiendo además preorientar a las macromoléculas antes de la colisión⁹⁹. Mostramos además que la unión E2c-ADN procede por dos rutas paralelas, tanto para la unión de oligonucleótidos con secuencia específica como para los inespecíficos (tabla 3.1). En base a esto, sugerimos que las moléculas de E2c se encuentran en más de una población conformacional *antes* de unir ADN. Estas poblaciones conformacionales pueden corresponderse con los *modos* de unión descritos en el mecanismo de unión al equilibrio. En este sentido, se observó que la formación del complejo 2:1 estaría dada por la unión de una de las poblaciones conformacionales al complejo 1:1, lo cual podría ocurrir luego de las etapas lentas de la unión (ver más abajo).

El efecto que las mutaciones puntuales ejercen sobre las etapas cinéticas de colisión fue analizado bajo dos hipótesis alternativas: o bien la mutación afecta la velocidad de colisión, o afecta el equilibrio de las poblaciones de conformeros (cap 3 y cap 5). Como ambas reacciones están limitadas por la difusión, ambas poblaciones conformacionales son diméricas y los efectos causados por las mutaciones afectan fuertemente la asociación cinética, nos inclinamos por la segunda interpretación. Atribuímos que las diferencias observadas sobre las k_{obs} de colisión se deben a cambios de las poblaciones relativas de cada conformero, lo cual indicaría que las mutaciones afectan la constante de equilibrio entre ambas poblaciones. Experimentos de estabilidad conformacional y cinética de desplegamiento de cada una de las mutantes permitirán aclarar este punto.

El efecto de la temperatura sobre la colisión nos permite asignar que la barrera de activación de la etapa de colisión es baja y mayoritariamente entrópica (tabla 3.3, Fig 3.12). Es también en esta etapa donde se detecta un desplazamiento de contraiones de la interfase de asociación (Fig 3.5), lo cual sugiere que los "complejos de encuentro" estarían estabilizados entrópicamente por la liberación de contraiones (Fig 3.12). Si esto fuera así, el efecto energético debería ser el mismo para la unión de E2c con cualquier oligonucleótido, independientemente de su secuencia de bases. Efectivamente, mostramos que la diferencia de energía libre de los "complejos de encuentro" de E2c con todos los oligonucleótidos ensayados es muy cercana (Fig 3.12). Además, dado que las velocidades de colisión están limitadas por la difusión, las barreras de activación de estos encuentros han de ser también similares.

Cabe destacar que el modelo sugerido define al "complejo de encuentro E2c-ADN" sobre una superficie energética isopotencial, independiente de la secuencia de bases del ADN. Esto es equivalente a decir que, una vez unida a ADN, es más favorable que E2c se *deslice* a lo largo de la doble hélice a que se *disocie* de la misma. Esto último se corresponde estrechamente con un mecanismo de reconocimiento sitio específico mediado por difusión facilitada conocido como "sliding" ¹⁰⁶. Se ha demostrado formalmente que la reducción en las dimensiones de búsqueda de 3D (colisión-disociación) a 1D (colisión-deslizamiento) acelera en órdenes de magnitud la probabilidad de encuentro de un sitio específico en el contexto de un genoma entero ⁶. Para el modelo de interacción LacR-ADN se ha estimado que, a temperatura ambiente, la proteína difunde a razón de 1000 Kpb/seg sobre una misma molécula de ADN ³⁷. ¹¹¹. Si esta velocidad es semejante para E2c (una proteína de tamaño comparable a LacR), a partir de las constantes cinéticas de disociación de ADN inespecífico podemos estimar que E2c se deslizaría por alrededor de 600 Kpb antes de disociarse de un duplex lineal.

REARREGLOS CONFORMACIONALES DEL COMPLEJO E2c-ADN

Se había propuesto con anterioridad que tanto E2c como el sitio específico de ADN sufren rearrreglos conformacionales al interactuar ^{37, 39, 40}. Es la distinción de cuáles son y cómo se dan estos rearrreglos la que permite una interpretación mecanística del proceso de reconocimiento secuencia-específico.

Para la unión de ADN específico, dos rearrreglos conformacionales son detectables por cambios en la fluorescencia de los Trp de E2c (Fig 3.4). El más lento de ellos coincide con la expulsión del fluoróforo ANS de la interfase de unión, sugiriendo que la desolvatación parcial de la interfase del complejo proteína-ADN ocurre en las últimas etapas de consolidación de la interacción. Por el contrario, ninguno de los oligonucleótidos de secuencia inespecífica fue capaz de desplazar el ANS unido a E2c indicando que, si bien la formación de complejos de colisión es igualmente rápida, estos no sufren rearrreglos conformacionales que incluyan la expulsión de solvente de la interfase. Esto refuerza la visión general de que la unión proteína-ADN inespecífico presenta un mayor grado de solvatación que los complejos con ADN específico (capítulo 1). Según se propuso para la interacción EcoRI-ADN, el agua secuestrada en las interfases puede proveer una fuente energética entrópica que sería requerida para la discriminación específica de secuencia ^{47, 48}.

El análisis termodinámico de las barreras de activación nos permitió discriminar las contribuciones energéticas de las cadenas laterales a las distintas etapas de unión. En principio, las barreras energéticas de los rearrreglos conformacionales son altas (~ 20 kcal/mol) respecto a la energía libre medida al equilibrio (~ -12kcal/mol), y al efecto producido por las mutaciones puntuales sobre éstas (~ 0,5 kcal/mol). Estas observaciones apoyan la inter-

pretación general de que la unión específica de ADN es el resultado de grandes valores absolutos de factores energéticos favorables y desfavorables ⁴⁶, en forma análoga a lo que ocurre en una reacción de plegamiento proteico ²⁵. Por otro lado, se ha visto que los procesos bioquímicos que requieren una alta capacidad de discriminación de sustratos similares (selección de aminoacyl-tRNAs por el ribosoma, de dNTPs por polimerasas, etc), utilizan estrategias de reconocimiento que involucran etapas separadas con altas barreras de activación ⁴⁶. Esto provee un mecanismo de "proofreading" cinético, donde las pequeñas diferencias entre los sustratos son *leídas* recursivamente en pasos sucesivos de reacción, lo cual da lugar a la ampliación de las pequeñas diferencias energéticas entre posibles sustratos (ya que al influir en varias etapas se multiplican) y permite lograr una alta capacidad de discriminación general ⁴⁶. Este tipo de mecanismos involucran *ajustes inducidos* ("induced-fit") en las macromoléculas reactantes que, para el caso E2c-ADN, podemos asignarlos a las barreras energéticas que separan los intermediarios de la unión. Tomados en conjunto, estas observaciones sugieren que la interacción proteína-ADN involucra cambios estructurales de plegamiento local acoplados al proceso de unión, y que éste está íntimamente ligado al proceso de reconocimiento secuencia-específico.

El conjunto de estados de transición (TSE) que separa el complejo de encuentro del siguiente intermediario es entrópicamente favorable ($\Delta S > 0$) y entálpicamente desfavorable ($\Delta H > 0$). Estos cambios son asignables a un evento de desplegamiento local, que produciría un aumento entrópico, junto con la redistribución de uniones no-covalentes. En línea con esto último, son las mutaciones sobre residuos que afectan la estabilidad de la hélice (A293, N294, F303) las que estabilizan el TSE de esta etapa cinética. Se distingue también un efecto de R300, R302 y de la capacidad de formar uniones de hidrógeno de C298 sobre el estado de transición. Esto sugiere que no solo el N-cap está afectando el estado de plegamiento local, sino que los cambios están asociados a múltiples residuos a lo largo de la hélice de reconocimiento. Proponemos entonces que los cambios conformacionales que ocurren en esta etapa son mayoritarios e involucran movimientos locales adjudicables al desplegamiento parcial de la hélice 1. Análisis estructurales de este Intermediario (por DC resuelto en el tiempo, doblemezcla con unión de fluoróforos, etc) permitirán evaluar esta hipótesis.

El segundo rearrreglo conformacional descrito tiene contribuciones mayoritariamente entálpicas (tabla 3.3, Fig 3.12), que atribuimos a la formación de los contactos cercanos en la interfase de unión. Para que estos contactos cercanos puedan tener lugar, las esferas de hidratación de las macromoléculas deben reestructurarse y es justamente en esta etapa donde se detecta la expulsión de ANS de la Interfase de unión. Los residuos que más afectan esta transición son F303, R300, K305, la carga positiva de K297, el carácter hidrofóbico de C298 y el anillo aromático de Y301. Dado que no contamos con el análisis termodinámico detallado del efecto de cada mutación, no nos es posible distinguir en esta instancia si los efectos de las mutaciones sobre la velocidad se deben a un cambio en el estado basal del reactante (el Intermediario 2) o a un efecto sobre el con-

junto de estados de transición que lo separan del estado final. De todas formas, podemos concluir que los efectos son asignables al menos a 7 de los 10 residuos expuestos de la hélice1, y particularmente al carácter hidrofóbico de 2 de ellos, apoyando la proposición de que es durante esta etapa donde se consolidan las interacciones de "lectura directa" de secuencia.

RUTAS DE ASOCIACIÓN PARALELAS. EFECTOS MICRO Y MACROSCÓPICOS.

Los experimentos de asociación-disociación por doble mezcla nos permitieron confirmar que los eventos cinéticos de primer orden son subsiguientes a la colisión inicial, asignables a los cambios conformacionales del complejo proteína:ADN recién discutidos, y que éstos se suceden en forma secuencial. Pudimos también medir en forma directa la disociación del "complejo de encuentro" y, quizá lo más sorprendente, observar que a los 10ms de iniciada la reacción de unión es posible detectar la formación del "complejo final". Esto constituye evidencia directa de la presencia de dos rutas macroscópicas de unión paralelas: una que procede por intermediarios con rearreglos lentos y otra de dos estados que permite la formación de un "complejo final" sin pasar por intermediarios definidos. Como la intensidad de la fluorescencia puede no estar relacionada con la concentración de las especies, no nos es posible asignar aquí qué porcentaje de la reacción procede por cada camino cinético. Sin embargo, en todos los ensayos de disociación de "complejos finales" se observaron dos eventos de disociación, indicando que la reacción E2c-ADN procede por dos rutas paralelas en ambos sentidos. El hecho de que la disociación de los "complejos finales" formados a tiempos cortos de asociación (10ms) muestren la mismas amplitudes relativas a lo largo de todo el tiempo de incubación, es indicativo de que las poblaciones de "complejos finales" están en rápido equilibrio. Esto sugiere además que estos complejos E2c-ADN están estructuralmente relacionados, y que intercambiarían en configuraciones geométricas separadas por barreras de activación relativamente pequeñas.

Uno de los análisis estructurales que pueden efectuarse sobre los estados de transición es el llamado análisis de valores de ϕ (Φ). Este tipo de análisis fue inicialmente planteado como una herramienta para estudiar los TSE de reacciones de plegamiento proteico y están basados en una analogía con los valores- β de Brønsted, clásicos de la fisicoquímica orgánica²⁵. Para evaluar estructuralmente el estado de transición de una reacción de plegamiento de dos estados, se considera el efecto que una mutación puntual causa sobre el TSE, respecto al efecto que causa la misma mutación al equilibrio, definiéndose para cada reemplazo,

$$\Phi = \Delta\Delta G^\ddagger / \Delta\Delta G_{eq} \text{ (para un esquema energético, ver Fig 5.1)}$$

Según la interpretación original, un valor de $\Phi = 1$ indica que el residuo realiza una interacción nativa en el TSE, mientras que $\Phi = 0$ indica que el residuo no forma esa interacción en el TSE²⁵. Cuando se detectan valores de Φ entre 0 y 1 se dice que la interacción está presente pero *debilitada*.

El caso de E2c-ADN presenta una ruta macroscópica de dos estados, y contamos con las diferencias energéticas causadas por las mutaciones puntuales sobre la energía libre de unión al equilibrio y los efectos que las mismas sustituciones causan sobre el TSE. Pudimos calcular entonces los valores de Φ para este camino de unión para 15 mutaciones puntuales (Fig 6.1). Los valores obtenidos se extienden mucho más allá de los valores esperados, pudiendo ser menores que 0 y mayores que 1, lo cual no tiene sentido físico en la interpretación clásica de Φ . Lo que podría ocurrir es que el marco teórico desde donde se plantea este tipo de análisis (un paradigma de plegamiento proteico de dos estados) no sea válido para una asociación proteína-ADN. Definir a las interacciones en los TSE como "nativas" o "no-nativas" puede no ser aplicable en todos los casos.

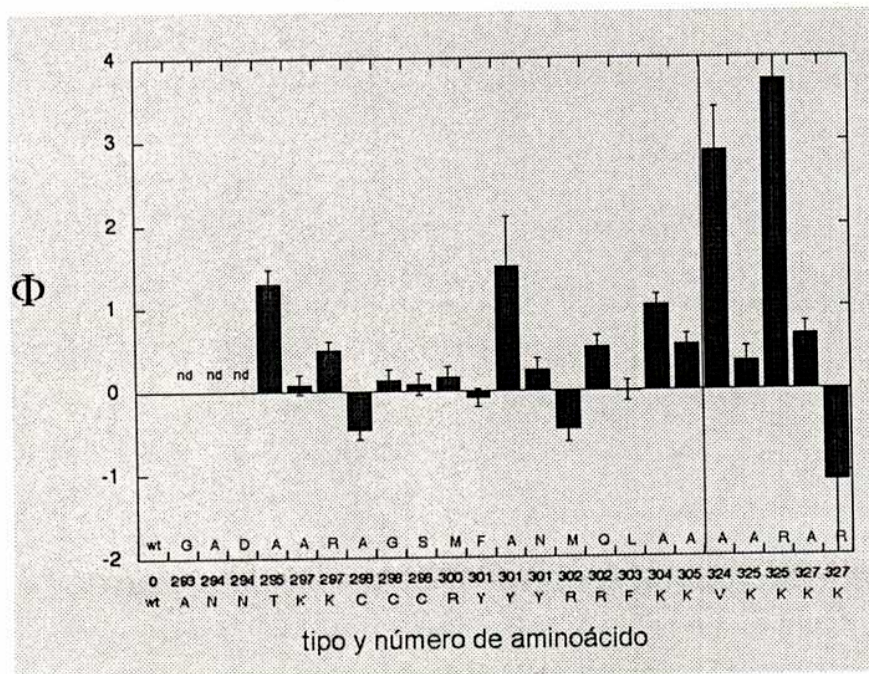


Figura 6.1: Valores de Φ de la unión E2c-ADN específico por la ruta de dos estados. Para cada mutante se calculó el valor de phi según el efecto que la mutación causa sobre la energía libre del TSE de dos estados, y el efecto que causa a la unión al equilibrio: $\Phi = \Delta\Delta G^\ddagger / \Delta\Delta G_{eq}$

Es muy probable que el rearreglo del TSE proceda por más de un camino microscópico y si así fuera, el efecto que una mutación puede causar sobre un TSE puede compensarse por la modificación de otras interacciones en el mismo estado de transición, sin que sea posible diferenciar las contribuciones individuales. Es decir, la modificación de un residuo podría afectar las interacciones de los otros residuos y no modificarse la velocidad de rearreglo. Esto es equivalente a decir que, a diferencia de lo que ocurre al equilibrio, los efectos causados por las mutaciones puntuales sobre los TSE no son independientes. La no-aditividad de las diferencias energéticas medidas para cada mutante en los TSE apoya indirectamente esta hipótesis.

Una interpretación alternativa de los valores de Φ fue propuesta recientemente el grupo de Ken Dill ⁷⁸. En este trabajo se muestra con un modelo minimalista exacto que los valores de Φ mayores a 1 o menores que 0 pueden deberse a la existencia de *caminos macroscópicos paralelos* de reacción. Según este modelo, las mutaciones influyen en Φ por afectar la población relativa de cada una de los caminos macroscópicos. Hemos presentado en esta tesis, por métodos independientes, que la asociación E2c-ADN procede efectivamente por dos rutas macroscópicas paralelas. Además, describimos que las mutaciones puntuales pueden afectar el equilibrio entre poblaciones conformacionales de E2c. Entonces, en este marco nos es posible asignar que los valores de Φ observados se corresponden con el modelo planteado inicialmente de caminos macroscópicos de unión paralelos. Si este esquema teórico se aplica a E2c-ADN, los valores Φ obtenidos sugieren que, además de proceder por rutas paralelas, podrían existir *cruzamientos* entre los caminos de unión ⁷⁸. Un análisis detallado de las rutas cinéticas de asociación de las proteínas con dobles-sustituciones, conjuntamente con las asignaciones termodinámicas de los TSE de las mismas, permitirá profundizar este aspecto del mecanismo de reconocimiento E2c-ADN.

ESTADOS CONFORMACIONALES DE E2C Y UNIÓN A ADN

Como mencionamos, una de las rutas de asociación entre E2c y ADN puede ser descrita como de dos estados. Dada la alta complejidad de la interacción, lo extenso de la interfase y la existencia de una ruta de unión con intermediarios separados por barreras energéticas altas, cabe preguntarse ¿cómo es posible que la asociación ocurra por una ruta rápida paralela?

Estudios de asociación entre macromoléculas, particularmente entre proteínas y ADN, proponen que el reconocimiento específico puede ser cinética y termodinámicamente más favorable si las zonas de unión a ADN están parcialmente desplegadas (son flexibles) y sufren procesos de plegamiento asociados a la unión ⁹¹. Cuando la proteína y el ADN se *ensamblan* pueden hacerlo gradualmente, pagando el costo entrópico a medida que se gana energía libre por la interacción y el plegamiento, es decir hay una correlación estrecha entre ambos componentes energéticos. Esto resulta análogo al paradigma propuesto para los paisajes energéticos ("landscapes") de reacciones de plegamiento proteico ^{28, 29}, y su reciente aplicación a procesos de interacción y reconocimiento específico entre macromoléculas ^{30, 101}. Un proceso cooperativo de unión-plegamiento de este tipo puede tener energías de activación muy bajas, donde el solvente puede expelerse gradualmente de la interfase, *lubricando* los movimientos necesarios para la formación de un complejo específico ⁹¹. Por el contrario, si la zona de unión tiene una estructura consolidada, existe un fuerte componente entrópico desfavorable a la unión, donde además se debe expulsar solvente para el anclado ("docking") específico. Estos dos últimos efectos resultan en barreras energéticas altas, coincidiendo con lo descrito para el camino de unión E2c-ADN que se sucede por dos intermediarios.

Ya hemos mencionado que existen evidencias independientes de que la hélice 1 de E2c no es una zona estructuralmente consolidada, y dadas las características de los caminos cinéticos y la existencia de dos poblaciones conformacionales de E2c, podríamos relacionar estas observaciones asignando a cada una de ellas ciertas diferencias estructurales. El conformero que procede por una asociación de múltiples estados es asignable al barril β -dimérico con las α -hélices consolidadas, como las presentadas por las resoluciones estructurales de la proteína libre, tanto en cristal como en solución. El otro estado conformacional, que se corresponde con la ruta de asociación de dos estados, presentaría mayores fluctuaciones conformacionales que el anterior, particularmente en las zonas de unión a ADN. Los intentos de enriquecer la población de E2c en este último estado resultaron en la formación de agregados insolubles, algo no del todo sorprendente si consideramos que sería a una población sin una estructura consolidada la que estamos tratando. Sin embargo hay evidencias indirectas que podemos considerar:

Si, como proponemos, existe una relación entre los *modos* de unión presentados por E2c al equilibrio y las poblaciones conformacionales mencionadas, es de esperar que la unión de dos dímeros de E2c por sitio de ADN este mediada por el conformero más flexible. Este pre-

sentaría menos restricciones conformacionales y podría converger en una interacción alternativa con el complejo 1:1 E2c-ADN, formando un complejo 2:1. Como se presentó en el capítulo 3, sólo una de las poblaciones conformacionales se une a un complejo 1:1, y resulta tener la misma velocidad de asociación que el confórmero de la ruta de dos estados, reforzando la proposición anterior. Una predicción de esta hipótesis es que, si fuera posible desplazar las poblaciones de confórmers de E2c al estado que presenta menores fluctuaciones (hélice consolidada), debería observarse un solo evento de unión (un único *modo* de unión) al mismo oligonucleótido. De acuerdo con esto último, en las condiciones de entorno en las cuales se resolvió la estructura consolidada en solución (pH 5.6), se observa un solo *modo* de unión a ADN (Fig 2.1). Por otro lado, experimentos recientes realizados con una preparación de E2c para la cual no se detectan cambios de DC ante un tratamiento térmico suave (indicando que no se desplaza la población de confórmers), presentan también un solo *modo* de unión al equilibrio.

En la figura 6.2 se esquematiza un modelo integrado del mecanismo de interacción E2c-ADN.

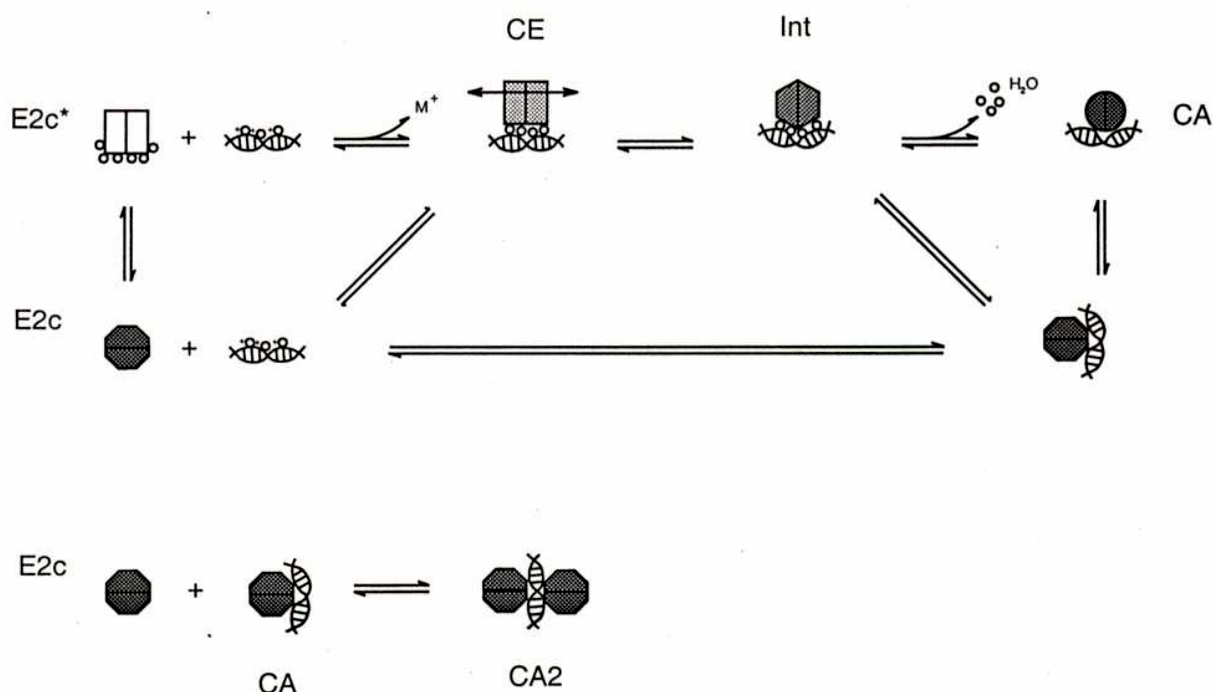


Figura 6.2: Modelo de interacción E2c-ADN. Arriba: unión de estequiometría 1:1 E2c:ADN: E2c existe en dos poblaciones conformacionales que unen ADN por rutas macroscópicas paralelas. Las flechas diagonales representan posibles cruzamientos entre ambas rutas. CE: complejo de encuentro, Int: intermediario 2, CA: complejo activo. Abajo: unión de estequiometría 2:1

MECANISMOS DE UNIÓN A ADN DE E2c EN UN MARCO BIOLÓGICO

Todos los estudios presentados en esta tesis fueron realizados *in vitro*, en condiciones mínimas de entorno, condiciones naturalmente distantes de las que podrían encontrarse en el entorno biológico de E2c, un núcleo celular. ¿Invalida esto las conclusiones biológicas que puedan derivarse del mecanismo de reconocimiento sitio-específico presentado? Sin intenciones de responder a esta pregunta, como tampoco de polemizar sobre un punto que excede los objetivos del presente trabajo y alcanza aspectos filosóficos profundos de la investigación científica actual, sólo mencionaremos que, hasta donde conocemos, los sistemas biológicos son extremadamente *robustos*. Es decir, las observaciones generales que se realizan para un sistema en una condición experimental particular son plausibles de aplicarse en otras condiciones, e incluso de extrapolarse entre sistemas, porque las características fundamentales de los mismos se conservan. Si se realizan con cautela y rigor científico, son éstas extrapolaciones las que permiten una comprensión integrada y general de un fenómeno particular.

Tomemos por caso la interacción específica proteína-sitio de ADN. Ha sido ampliamente descrito que la interacción específica es fuertemente dependiente de factores físico-químicos del entorno, como la fuerza iónica o la temperatura. Sin embargo, estos mismos factores afectan en grados comparables la unión con ADN inespecífico. Como la unión biológicamente relevante depende de la relación entre la unión específica e inespecífica, la ocupancia del sitio *operador* resulta ser independiente de la temperatura y de la fuerza iónica del entorno celular ⁵⁷. En algunos casos, incluso los valores absolutos de las constantes de afinidad medidas *in vitro* resultan estar en los rangos estimados como fisiológicamente relevantes, lo cual implica que la interacción bioquímica también es robusta ⁵⁷. Basándonos en lo descrito a lo largo de toda esta tesis, arriesgaremos dos breves hipótesis funcionales sobre el posible mecanismo de interacción E2c-ADN *in vivo*.

Hemos descrito que E2c puede presentarse en más de una población conformacional, que presentan notables diferencias en cuanto a sus características de unión de ADN. Como para muchos otros factores de transcripción, empezando por el prototípico LacR, la ocurrencia de estados conformacionales puede corresponder a la existencia de un "switch" de regulación de la actividad proteica. Describimos aquí que la temperatura, el pH o las mutaciones puntuales pueden alterar el equilibrio de las poblaciones de E2c, pero es poco probable que estos factores sean responsables de la regulación de la actividad *in vivo*. Resulta biológicamente más atractivo suponer que esta fluctuación conformacional sea regulada por otros eventos, muy probablemente mediados por la interacción de E2c con otras proteínas. Se ha descrito que una amplia variedad proteica interactúa con E2c *in vivo* y que esto regula a su vez la actividad biológica de E2c (capítulo 1). Algunas de estas proteínas son candidatos naturales de operar sobre el "switch" bioquímico de E2c que, proponemos, puede estar basado en el *bisagra* conformacional propuesta sobre la hélice 1. Por otro lado, sería muy interesante evaluar el efecto fisiológico de la virtual anulación del "switch", causado por las mutaciones en el N-cap de la hélice 1, que desplazan el equilibrio de la población de conformeros hacia un estado mayori-

tario. Estas mutantes podrían presentar diferencias en las diferentes actividades biológicas de E2c (represión/activación de la transcripción, iniciación de la replicación, segregación episomal, etc), que podrían brindar indicios sobre la funcionalidad de estos cambios conformacionales *in vivo*.

Finalmente, una última extrapolación del mecanismo cinético de reconocimiento sitio-específico. En primera instancia, vamos a suponer que el mecanismo de unión por dos rutas macroscópicas paralelas se conserva *in vivo*, pero que los valores de las constantes cinéticas pueden no hacerlo. Si esto así fuera, es posible plantear que una de las rutas corresponde a la asociación y la otra a la disociación de E2c del sitio específico. Basados en analogías simples con modelos de reconocimiento mediados por "proofreading" cinético, es de esperar que la ruta de asociación corresponda al camino macroscópico con intermediarios y la de disociación al de dos estados. En este caso, las altas barreras energéticas que separan los intermediarios de unión pueden ser planteadas como "checkpoints" del proceso de reconocimiento. Esto nos permite proponer un esquema informático¹⁰ del reconocimiento sitio específico (Fig 6.3).

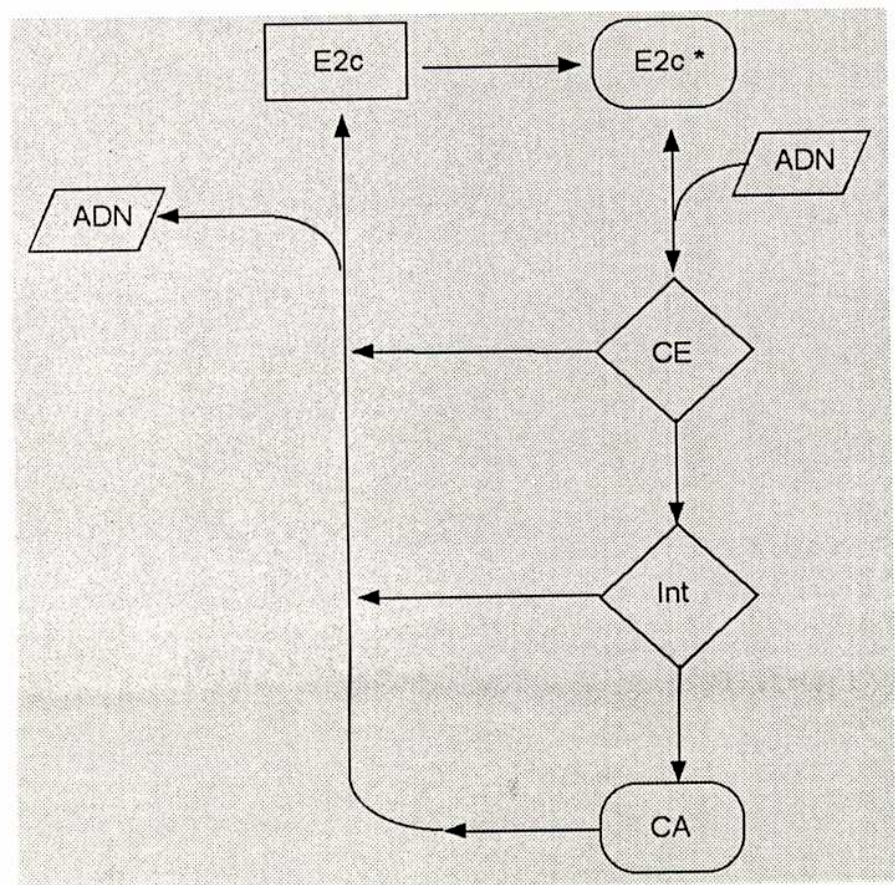


Figura 6.3: Modelo informático de discriminación de sitios de ADN. E2c se encuentra en dos conformaciones, una de las cuales une ADN y forma el complejo de encuentro (CE). La formación del complejo activo (CA) está mediada por dos "checkpoints" de reconocimiento de secuencia. La disociación de CA procede por una ruta paralela.

La unión se inicia desde la proteína libre en la población conformacional de la estructura consolidada. La proteína se une a ADN de manera inespecífica y difunde en forma unidimensional, con una procesividad que podemos estimar en al menos el tamaño del genoma de HPV (8 Kpb). Dependiendo de las afinidades relativas por las diferentes secuencias, E2c pasará más tiempo en unas regiones que en otras, pero sólo algunos de los sitios visitados permitirán que se transpase la primer barrera de alta energía. Para pasar este "checkpoint", el sitio de ADN debe permitir un rearrreglo conformacional en E2c que presumiblemente está asociado a un cambio conformacional de la hélice-1. Formado el siguiente intermediario de reacción, existe otra alta barrera energética a superar que involucra la expulsión de solvente de la interfase y el contacto directo de los motivos específicos de secuencia. Si alguna de estas barreras no es superada en un tiempo límite (centrado en la vida media de los intermediarios), el proceso de unión es interrumpido por el *desvío* hacia la ruta de disociación. Esta *bifurcación* puede corresponderse con la sugerencia, desprendida de los valores de Φ , de que los caminos macroscópicos paralelos pueden *cruzarse*. En cambio, si la secuencia de ADN permite superar ambos "checkpoints", se da lugar a la formación de un complejo consolidado que presenta una alta estabilidad y alta vida media: el complejo activo en el sitio específico. El ciclo se cierra cuando E2c se despegue del sitio de unión, cosa que sería más favorable si se realice por la ruta macroscópica que caracterizamos como de "dos estados". Según los correlatos estructurales planteados, para despegar a E2c del sitio específico sería suficiente con *flexibilizar* la hélice de reconocimiento para que se inicie la ruta de disociación. Esto bien puede estar modulado por el contacto de E2c y otras proteínas, particularmente en la zona del N-cap de la hélice 1. En relación con esto último se ha reportado que E2 de HPV16 es *activamente* desplazada del sitio de unión por la proteína E1⁸⁸, con la cual se han reportado también interacciones mediadas por dominio C-terminal (E2c)⁷⁴. Si bien este modelo general resulta hoy altamente especulativo, las predicciones que se desprenden son empíricamente contrastables y resultan ser biológicamente relevantes a muchos sistemas de interacción proteína-ADN.

Todo lo discutido durante este trabajo podría converger en la idea central de que las interacciones proteína-ADN incluyen una amplia variedad de procesos bioquímicos acoplados. Confiamos en que una interpretación biofísica de estos fenómenos puede contribuir significativamente a la comprensión de los mecanismos moleculares que subyacen el universo biológico cotidiano. Utilizando un sistema relativamente simple, pudimos empezar a responder preguntas energéticas y mecánicas básicas, que resultan de difícil aproximación en modelos más complicados. Esperamos con esto contribuir a que en el futuro cercano contemos con nuevas herramientas teóricas y experimentales que permitan abordar en profundidad los mecanismos moleculares de una reacción biológica tan relevante como lo es el reconocimiento sitio-específico de una secuencia de ADN.

MATERIALES Y MÉTODOS

"Lejos de ser los Rimbaud del microscopio que un público sentimental gusta imaginarse, los investigadores en biología molecular no son más que técnicos honrados, carentes de genio, que leen "Le Nouvelle Observateur" y sueñan con irse de vacaciones a Groenlandia"

Michelle Houellebecq

REACTIVOS GENERALES

Todos los reactivos utilizados para las mediciones espectroscópicas fueron del más alto grado de pureza disponible. Tampón BisTris, Urea, Cloruro de Guanidina, Cloruro de Sodio, Sales de Fosfato y DTT fueron adquiridos comercialmente de ICN (Costa Mesa, CA, EEUU). Isotiocianato de Fluoresceína, ANS y Bis-ANS de Molecular Probes (Eugene, OR, EEUU). Las enzimas de restricción, ligasa, y polinucleótido quinasa son de New England Biolabs (Beverly, MA, EEUU), y la *Pfu* DNA polimerasa de Stratagene (La Jolla, CA, EEUU). Los demás reactivos de Sigma-Aldrich (St. Louis, MO, EEUU).

OLIGONUCLEÓTIDOS SINTÉTICOS

Los oligonucleótidos de simple cadena fueron adquiridos comercialmente de IDT (Coralville, IA, EEUU). Cuando se indica FITC (o FI) un motivo de isotiocianato de fluoresceína fue introducido durante la síntesis en el extremo 5' del oligo mediante un conector de 6 carbonos. Luego de la síntesis, los oligonucleótidos fueron purificados por geles de poliacrilamida.

Los oligonucleótidos de doble cadena utilizados fueron preparados mezclando cantidades equimolares de las cadenas complementarias, entre 10 y 30 μ M para los oligonucleótidos modificados con FITC, 100 μ M para los no-modificados. La concentración de los oligonucleótidos se determinó espectrofotométricamente. Las reacciones de apareamiento de hebras se realizaron en 10mM BisTris-HCl pH7.0, 20mM NaCl, incubados por 5 min a 95°C y enfriados lentamente hasta 25°C durante 16 horas. Este procedimiento rinde oligonucleótidos de doble cadena sin cantidades detectables de simple cadena, a juzgar por corridas en geles nativos de poliacrilamida (no mostrado).

las secuencias de los oligonucleótidos utilizados son:

s35 A: 5' - GTA ACCG AAAT CGGT TGA - 3'
 s35 B: 5' - TCA ACCG ATTT CGGT TAC - 3'
 sBPV: 5' - CCG ACCG ACGT CGGT CGG - 3'
 EBNA: 5' - GGG TAG CAT ATG CTA CCC - 3'
 Iset A: 5' - ACA TGG ACC TGT CAA GTA - 3'
 Iset B: 5' - TAC TTG ACA GGT CCA TGT - 3'

(subrayado, la secuencia consenso de E2)

Los demás oligonucleótidos sintéticos, utilizados para las construcciones plasmídicas y las mutantes sintetizadas, fueron adquiridos a la misma empresa. Para estos, no se encargó purificación alguna más allá de la estándar.

EXPRESIÓN RECOMBINANTE Y PURIFICACIÓN DE E2C

HPV16-E2c (hE2c)

La expresión y purificación de hE2c se llevó a cabo según el protocolo descrito por Mok *et. al.* ⁷³. No tan brevemente: bacterias *Escherichia coli* BL21 (DE3) portando el plásmido pTZ18u:E2c fueron inocuadas (al 2% con un cultivo saturado) en erlenmeyers conteniendo 600ml de medio 2TY (16 g peptona, 10g extracto de levadura, 5g NaCl /litro) a 37°C con 100 µg/ml de ampicilina. Cuando el cultivo alcanza la fase exponencial ($DO_{600} = 0,5$) se agregó IPTG hasta 0,4 mM y se incubó por 16 hs a 37°C, con agitación orbital a 300 RPM. Las bacterias se cosecharon por centrifugación a 2000g, por 30 min, y se resuspendieron en 0,1 vol de buffer de lisis (100 mM Tris-HCl pH 6.8, 0,6M NaCl, 1mM EDTA, 10mM 2-mercaptoetanol). La suspensión fue sonicada dos veces a 0°C, y centrifugada a 3000g por 20 min. Al sobrenadante, que contiene la fracción soluble de hE2c, se le agregó sulfato de amonio al 60% de saturación (361 g/l) y se incubó a 4°C con agitación por 16hs. El precipitado se colectó por centrifugación a 12000g por 15 min, y se resuspendió en buffer A (100mM Tris-HCl pH 6.8, 0,6M NaCl, 2mM 2-mercaptoetanol). La fracción soluble se sembró en una columna de heparina (BioSeptra, Villeneuve la Garenne, Francia) previamente equilibrada en 50mM Tris-HCl pH 8.0, 0,6M NaCl, 2mM 2-mercaptoetanol. La columna fué lavada con al menos 5 vol del mismo buffer y eluida con un gradiente iónico lineal hasta 2M NaCl. Las fracciones conteniendo hE2c >85% pura (entre 0,9 y 1,1 M NaCl) fueron juntadas y dializadas contra 2 cambios de 50 vol de buffer F (10mM Acetato de Sodio pH 5,6, 0,2 M NaCl, 1mM DTT) y concentradas por ultrafiltración en Centriprep-3 (Amicon, Bedford, MA, EEUU). La fracción soluble fué filtrada por un tamiz molecular Superdex75 (Pharmacia Biotech, Uppsala, Suecia), donde hE2c eluye como un único pico de 18-20 KDa. El procedimiento detallado rinde alrededor de 40 mg de hE2c >97% pura por litro de cultivo inicial. La proteína purificada fue guardada a -70°C luego de congelarla rápidamente con N₂ líquido.

Comentario acerca de la actividad de unión a ADN de hE2c

El hecho que la proteína eluya con alta fuerza iónica de la columna de heparina sugiere que la actividad de unión a ADN es alta. Asimismo, hay numerosas evidencias de que el procedimiento descrito rinde hE2c dimérica, plegada, con alta actividad de unión, como la utilizada en ensayos cristalográficos. De todas formas realizamos un ensayo para estimar la actividad de unión a ADN, en donde incubamos la preparación de hE2c a 10µg/ml con un exceso 2 veces molar de un oligonucleótido doble cadena s35, modificado en 5' con biotina. Los complejos formados fueron incubados con "bolitas" magnéticas de estreptavidina (Promega, Madison, WI, EEUU) y luego removidos de la solución utilizando un imán. La concentración de proteína en el sobrenadante fue igual a la de un control en el cual no se agregó hE2c, indicando que la preparación de proteína es >97% activa en tanto a unión de ADN específico.

BPV1-E2c (bE2c)

La expresión de bE2c se llevó a cabo en forma recombinante utilizando bacterias *Escherichia Coli* JM109 portando el plásmido pTZE2kz ⁶⁹. Las células se crecieron en medio 2TY, 100µg/ml ampicilina, a 37°C hasta DO₆₀₀ = 0,5. Se agregó IPTG a 0,4mM y se infectaron con el fago M13/T7 (Invitrogen, SanDiego, CA, EEUU) a una multiplicidad de infección de 5. Se incubaron por 16hs a 37°C, se colectaron por centrifugación, se resuspendieron en buffer de lisis y se sonicaron 2 veces a 0°C. Los cuerpos de inclusión se colectaron centrifugando a 15000g por 20 min, se lavaron con 8M urea, y se disolvieron en 500mM Tris-HCl pH6.8, 10mM DTT y 6M GuHCl. El replegado se realizó dializando contra 100 vol de 50mM Tris-HCl pH=6.8, 0,15M NaCl y 2 mM 2-mercaptoetanol. La fracción soluble se sembró en una columna de Heparina y se continuo la purificación como se describió para hE2c, pasado la muestra por un tamiz molecular. El procedimiento rinde alrededor de 2 mg/l de bE2c 98% pura.

Conjunto de mutantes sitiodirigidas de hE2c

Todas las mutantes descritas se expresaron en forma soluble en *E. coli* BL21 (DE3) y se purificaron siguiendo el procedimiento descrito para la proteína salvaje, con algunas excepciones, a saber: la fuerza iónica del buffer A se redujo a 0,2M NaCl y se eluyó con un gradiente lineal desde 0,2 a 2M NaCl. Los rendimientos totales de proteína variaron de 1 a 40 mg/l, y en todos los casos la pureza obtenida, estimada por corridas en PAGE es mayor al 98%.

DETERMINACIÓN DE LA CONCENTRACIÓN DE PROTEÍNA

Cuando se trata de extractos o proteínas impuras, se utilizó el método de Bradford ².

Para todas las proteínas purificadas utilizadas se determinó su concentración espectrofotométricamente utilizando el coeficiente de absorción calculado según ⁶¹:

$$A_{280} \text{ (mM}^{-1}, \text{cm}^{-1}\text{)} = 5,5 \text{ \#W} + 1,49 \text{ \#Y}$$

donde #W y #Y son el número de residuos de Trp y Tyr, respectivamente.

El método fue validado con la proteína hE2c salvaje midiendo la concentración por el método de Bradford.

CONSTRUCCIÓN DE LAS MUTANTES SITIODIRIGIDAS

Las mutaciones sitiodirigidas se realizaron utilizando el método de PCR inversa ¹⁶. El mismo consiste en utilizar oligonucleótidos sintéticos, que incluyan la mutación deseada, y amplificar todo el plásmido por PCR. Como las mutaciones diseñadas están próximas en la secuencia nucleotídica, se diseñaron juegos de cebadores conteniendo las mutaciones, tal que aparezcan en una misma hebra. Estos se utilizaron en combinación con un único cebador salvaje, que aparece en la complementaria del plásmido pTZ18u:E2c utilizado como molde.

Previamente a la PCR, los cebadores fueron fosforilados en 5' utilizando polinucleótido quinasa y ATP, según los instructivos del fabricante. Para la síntesis de ADN se utilizó *Pfu* polimerasa, dado que esta presenta alta fidelidad y no se detecta actividad de transferasa terminal, típica de la polimerasa Taq. La PCR se realizó por 25 ciclos rindiendo, en general, un único producto de alrededor de 3,5 Kpb. El molde de reacción fue degradado selectivamente utilizando la enzima de restricción *DpnI*. Esta enzima corta solamente sitios GATC metilados, por tanto no degrada el producto de amplificación. La ligación de los extremos romos del plásmido se realizó utilizando la ligasa del fago T4 por 16hs entre 16°C y 25°C.

Las mezclas de ligación se utilizaron para transformar *E. coli* DH5 α , utilizando el método de shock térmico en cloruro de calcio ². Se replicaron entre 6 y 10 colonias al azar y se realizaron cultivos de 2ml en medio LB. De éstos se preparó ADN plasmídico según el protocolo de lisis alcalina ² y se secuenciaron entre 400 y 700 pb, utilizando equipos de secuenciación automática ALF (Pharmacia Biotech, Uppsala, Suecia). Se tuvo especial cuidado en obtener la secuencia de todo el marco de lectura de E2c, y se descartaron los clones que contenían mutaciones más allá de las buscadas. Según las secuencias analizadas, me permito estimar que alrededor de 1/7 de los clones obtenidos de cada reacción contienen sólo la mutación diseñada.

ENSAYO DE RETARDO DE LA MOVILIDAD ELECTROFORÉTICA

El ensayo de EMSA o "gel shift" se llevó a cabo en condiciones de baja fuerza iónica ². Las mezclas de reacción fueron incubadas en las condiciones estándar de unión, por 30 min a 25°C en un volumen final de 10 μ l. Las mezclas se sembraron en un gel de poliacrilamida al 6% conteniendo 2,5% de glicerol, 25mM Tris-HCl pH 8.3, 190mM Glicina, 1mM DTT. El gel se corrió refrigerado a 4 V/cm durante 2 horas. Las bandas de fluoresceína se detectaron por transiluminación UV. Después de la visualización y la documentación, el gel fue fijado y teñido con Ag, a fin de revelar las bandas proteicas.

DICROÍSMO CIRCULAR

Los espectros de dicroísmo circular fueron registrados en un espectropolarímetro Jasco J-810 (Nikota, Japón). Entre cinco y diez espectros fueron promediados para cada muestra. Cuando no se indica lo contrario, la temperatura se mantuvo constante a 25 ± 0.1 °C con un controlador peltier. Para los espectros de proteínas en la región de UV lejano, se calculó la elipticidad media por residuo según:

$$MRW = mdeg / 10 \cdot L \cdot [M] \cdot (\#p)$$

donde mdeg es la señal medida, L el paso de luz (en cm), [M] la concentración de la muestra (en molar), #p el número de enlaces peptídicos de la proteína en cuestión. Como para E2c consideramos que trabajamos siempre con el dímero, todas las concentraciones se expresan de dímeros, así como la cantidad de enlaces peptídicos.

Los espectros de CD de ADN se registraron entre 320 y 250 nm. La contribución de las proteínas se sustrajo directamente a partir de espectros de la proteína sola. Es de esperar que la contribución de E2c en esta zona del espectro no cambie significativamente, por lo que asumimos que evaluamos solo los cambios dicróicos del ADN³⁰. Los espectros diferenciales fueron obtenidos restando al espectro del complejo el de la proteína libre y el del ADN libre.

ENSAYOS DE UNIÓN A ADN EN SOLUCIÓN EQUILIBRIO

Todos los ensayos de unión a ADN se realizaron con un espectrofluorímetro Aminco Bowman Series 2, arreglado en geometría "L" (Spectronic Co. Rochester, NY, EEUU). Cuando el fluoróforo fué FITC, se excitó a 490 nm y se colectó la emisión a 520nm, con 4 nm ancho de banda, a menos que se indique lo contrario. Cuando la concentración de fluoresceína fué menor que 20nM, se colectó la fluorescencia total utilizando un filtro de corte bajo 520nm (Schott, Mainz, Alemania).

Todas las titulaciones se llevaron a cabo agregando pequeñas cantidades del ligando variable (entre 2 y 20 µl) a una cubeta conteniendo el ligando fijo. Se homogeneizó la solución con pipeta automática, succionando y desplazando al menos medio volumen total, 5 veces. Se dejó luego equilibrar la muestra por dos minutos y se tomó la medición correspondiente. En ningún caso la dilución del ligando fijo superó el 15% del volumen total. Los datos de intensidad de fluorescencia se corrigieron linealmente según la dilución.

Las colecciones anisotrópicas se realizaron utilizando polarizadores lineales y colectando las intensidades verticales y horizontales a las mismas longitudes de onda antes mencionadas. Se promediaron al menos 5 mediciones a 0° y 90° y se calculó la anisotropía (r) según ⁵²:

$$r = (I_{vv} - G \cdot I_{vh}) / (I_{vv} + 2G \cdot I_{vh})$$

donde I es la intensidad de fluorescencia, los sufijos V y H refieren a Vertical y Horizontal, primero excitación y luego emisión, y el factor de corrección G está definido como:

$$G = I_{hv} / I_{hh}$$

Dado el pequeño cambio de rendimiento cuántico del motivo FITC asociado a la unión de ADN, no fue necesario realizar correcciones de intensidad sobre los valores anisotrópicos ⁵².

Las condiciones estándar del ensayo fueron fijadas a: 25mM BisTris-HCl pH 7.0, 200 mM NaCl y 1mM DTT. En todos los casos la temperatura se mantuvo constante a 25 ± 0.1 °C utilizando un baño térmico acoplado al fluorímetro.

CINÉTICA

El curso cinético de las reacciones se midió utilizando un aparato de flujo detenido (stopped-flow) Applied Photophysics SX18.MV (Leatherhead, Reino Unido). Para este equipo el tiempo muerto de reacción es de 2ms y permite la colección de 1000 puntos en intervalos de tiempo variables entre 0-20ms hasta 0-1000 sec. Todas las concentraciones reportadas corresponden a aquellas resultantes de la mezcla de volúmenes iguales de ambas jeringas. Las reacciones fueron monitoreadas siguiendo la fluorescencia de FITC o de Trp, excitando a 490 nm o 280 nm, respectivamente. La emisión se colectó a 90° con filtros de corte bajo de 515 o 320nm (Schott, Mainz, Alemania). Cada medición reportada corresponde al promedio de entre cinco y diez colecciones de 1000 puntos cada una. Los datos se ajustaron a las ecuaciones exponenciales descriptas más abajo por minimización no-lineal de cuadrados, utilizando el software provisto por el fabricante. En todos los casos la temperatura se mantuvo constante a 25 ± 0.1 °C utilizando un baño térmico, a menos que se indique lo contrario.

Los experimentos de asociación - disociación por doble-mezcla se realizaron mezclando 0,5 y 0,5 volúmenes de proteína y FI-ADN. La colección de datos se inició luego del "tiempo de incubación", cuando se mezcla la muestra con un volumen más de ADN sin marca fluorescente. El tiempo preciso transcurrido entre ambas mezclas se recupera por medición directa de cada disparo, a través de sensores dispuestos en el aparato a tal fin.

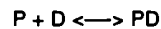
MUTANTES HE2C

La medición de la interacción con ADN de cada una de las mutantes de E2c descritas se realizó siguiendo el mismo procedimiento descrito para la proteína salvaje. A fin de homogeneizar las preparaciones y los ensayos, cada proteína fue descongelada, redializada por una noche, clarificada por centrifugación y re-cuantificada el día anterior a realizar las mediciones de Interacción con ADN. Se realizaron entre 3 y 5 Isotermas a diferentes concentraciones de fijas de ADN. Durante los días en los cuales se realizaron las mediciones, las muestras fueron mantenidas a 0°C. Como control interno, se siguieron los mismos procedimientos con la proteína salvaje, arrojando en todos los casos mediciones energéticas consistentes.

↓

MODELOS MATEMÁTICOS UTILIZADOS Y AJUSTES A LOS MISMOS EQUILIBRIO

Las isotermas de unión proteína:ADN medidas al equilibrio fueron ajustadas a un modelo de interacción:



donde P es [proteína], y D [ADN], y la concentración total de ambas es

$$P_t = P + PD$$

$$D_t = D + PD$$

la constante de disociación (K_D) se define como

$$K_D = [P] [D] / [PD]$$

de lo cual se deduce que la concentración de complejo [PD] para cada punto será:

$$[PD] = 0,5 ([D]+[P]+K_D) - ([D]+[P]+K_D)^2 - (4 [D]+[P])^{0,5} \quad (\text{ecuación 7.1})$$

asi, sabiendo la concentración total [P_t] y [D_t] en cada punto, y dado que la señal fluorescente medida es lineal con la concentración de [PD], del ajuste de una isoterma a la ecuación descripta se obtiene en valor de K_D .

En algunos casos las isotermas mostraron desviaciones lineales en la señal fluorescente más allá de la señal reportada por la unión. En esos casos, se sumó un término lineal al ajuste: $(\pm) m1 \cdot [P]$.

La calidad del ajuste se midió por la desviación de las diferencias de cuadrados y por desvío aleatorio de residuales.

CINÉTICA

Las curvas de asociación/disociación cinéticas se ajustaron a modelos de decaimiento exponencial de la forma:

$$F(t) = A \cdot \exp(k \cdot t) + C$$

donde $F(t)$ es la fluorescencia al tiempo t , A la amplitud total del cambio, C la fluorescencia a $t = \infty$, y k la constante cinética de reacción. Cuando una sola exponencial no fue suficiente para un buen ajuste a los datos, se sumaron términos exponenciales de la misma forma:

$$F(t) = A1 \cdot \exp(k1 \cdot t) + A2 \cdot \exp(k2 \cdot t) + C$$

donde $A1$ y $A2$ son las amplitudes de las fases, $k1$ y $k2$ las constantes de reacción respectivas. En ningún caso se utilizaron más de 3 exponenciales en un mismo ajuste. Para el ajuste de los datos se utilizó un algoritmo de aproximación por cuadrados mínimos no lineales, provisto por el fabricante del equipo.

Las curvas de pseudoprimer orden responden a la forma:

$$k_{obs} = k_{on} \cdot [X] + k_{off}$$

donde k_{obs} es la constante observada, k_{on} la constante de asociación de segundo orden, $[X]$ la concentración de ligando variable, y k_{off} la constante de disociación.

PARÁMETROS TERMODINÁMICOS

De la dependencia de las k_{obs} con la temperatura se calcularon los parámetros termodinámicos de activación asociados a los conjuntos de estados de transición (TSE), según lo descrito por Jackson y Fersht ⁴⁴. En todos los casos se realizaron gráficos de Eyring ($\ln k_{obs} / T$ vs. $1 / T$) y se ajustaron por cuadrados mínimos no lineales a las siguientes ecuaciones:

en caso de dependencia lineal:

$$\ln (k_{obs}/T) = \ln k_B/h + \Delta S^\ddagger/R - \Delta H^\ddagger/R \cdot 1/T \quad (\text{ecuación 7.2})$$

donde T es temperatura, k_B la constante de Boltzmann, h la constante de Planck, R la con-

stante de los gases, $\Delta H^\ddagger$ la entalpía de activación y $\Delta S^\ddagger$ la entropía de activación. Estos dos últimos están relacionados a la energía libre de activación según:

$$\Delta G^\ddagger = \Delta H^\ddagger - T\Delta S^\ddagger$$

Cuando se detectaron desviaciones de la linealidad de un gráfico de Eyring, se siguió el análisis presentado por Chen y Matthews ¹³ ajustando los datos a:

$$\ln (k_{obs}/T) = A + B (T_0/T) + C \ln (T_0/T) \quad (\text{ecuación 7.3})$$

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

$$B = (\Delta C_p^\ddagger - \Delta S^\ddagger(T_0))/R - \Delta G^\ddagger(T_0)/R T_0$$

$$C = -\Delta C_p^\ddagger / R$$

donde T_0 es la temperatura de referencia (298,14 °K), y $\Delta C_p^\ddagger$ la capacidad calorífica de activación, definida como:

$$\Delta C_p^\ddagger = \partial \Delta H^\ddagger / \partial T$$

SOFTWARE

Todos las gráficas de puntos y los ajustes a los datos se realizaron con el programa KaleidaGraph (Albeck Software). Las representaciones estructurales de proteínas y ADN se realizaron con el programa SwissPDBViewer ³¹, según las coordenadas depositadas en el ProteinDataBank ³². Para los cálculos de superficies se utilizó el programa SurfaceRacer ¹⁰², utilizando como sonda una esfera de 1.4 Å de radio.

Todos los programas utilizados se ejecutaron en una computadora Apple Macintosh G3 corriendo a 500 Mhz sobre el sistema operativo MacOSX , .

REFERENCIAS

"Alles Gescheide ist schön gedacht worden, man muss nur versuchen
es noch einmal zu denken"

Johann Wolfgang von Goethe

- 1) Alonso, L. G., Garcia-Alai, M. M., Nadra, A. D., Lapena, A. N., Almeida, F. L., Gualfetti, P. & Prat-Gay, G. D. (2002). High-risk (HPV16) human papillomavirus E7 oncoprotein is highly stable and extended, with conformational transitions that could explain its multiple cellular binding partners. *Biochemistry* 41(33), 10510-8.
- 2) Ausbel, F., Brent, R., Kingston, R. E., Moore, D. D., Siedman, J. G., Smith, J. G. & Struhl, K. (1997). *Short protocols in molecular biology*, John Wiley & Sons, Inc, New York.
- 3) Avery, O. T., McLeod, C.M. & McCarthy M. (1944). Induction of transformation by a desoxyribonucleic acid fraction isolated from *Pneumococcus* Type III. *J. Exp. Med.* 79, 137-158.
- 4) Bastien, N. & McBride, A. A. (2000). Interaction of the papillomavirus E2 protein with mitotic chromosomes. *Virology* 270(1), 124-34.
- 5) Bedrosian, C. L. & Bastia, D. (1990). The DNA-binding domain of HPV-16 E2 protein Interaction with the viral enhancer: protein-induced DNA bending and role of the nonconserved core sequence in binding site affinity. *Virology* 174(2), 557-75.
- 6) Berg, O. G., Winter, R. B. & von Hippel, P. H. (1981). Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. *Biochemistry* 20(24), 6929-48.
- 7) Bochkarev, A., Barwell, J. A., Pfuetzner, R. A., Bochkareva, E., Frappier, L. & Edwards, A. M. (1996). Crystal structure of the DNA-binding domain of the Epstein-Barr virus origin-binding protein EBNA1, bound to DNA. *Cell* 84, 791-800.
- 8) Bogan, A. A. & Thorn, K. S. (1998). Anatomy of hot spots in protein interfaces. *J Mol Biol* 280(1), 1-9.
- 9) Boner, W. & Morgan, I. M. (2002). Novel cellular interacting partners of the human papillomavirus 16 transcription/replication factor E2. *Virus Res* 90(1-2), 113-8.
- 10) Bray, D. (1995). Protein molecules as computational elements in living cells. *Nature* 376(6538), 307-12.
- 11) Bray, D. (2003). Molecular networks: the top-down view. *Science* 301(5641), 1864-5.
- 12) Brooijmans, N., Sharp, K. A. & Kuntz, I. D. (2002). Stability of macromolecular complexes. *Proteins* 48(4), 645-53.

- 13) Chen, X. & Matthews, C. R. (1994). Thermodynamic properties of the transition state for the rate-limiting step in the folding of the alpha subunit of tryptophan synthase. *Biochemistry* **33**(20), 6356-62.
- 14) Church, G. M., Sussman, J. L. & Kim, S. H. (1977). Secondary structural complementarity between DNA and proteins. *Proc Natl Acad Sci U S A* **74**(4), 1458-62.
- 15) Ciuffo, G. (1907). Innesso positivo con filtrato di verruca volgare. *Giorn Ital Mal Venereol* **48**, 12-17.
- 16) Clackson, T., Dettlef, G. & Jones, P. (1991). PCR, a Practical Approach. In *PCR, a Practical Approach* (McPherson, M., Quirke, P. & Taylor, G., eds.), pp. 202. IRL Press, Oxford.
- 17) deHaseth, P. L., Lohman, T. M. & Record, M. T., Jr. (1977). Nonspecific interaction of lac repressor with DNA: an association reaction driven by counterion release. *Biochemistry* **16**(22), 4783-90.
- 18) DeLano, W. L. (2002). Unraveling hot spots in binding interfaces: progress and challenges. *Curr Opin Struct Biol* **12**(1), 14-20.
- 19) Dell, G. & Gaston, K. (2001). Human papillomaviruses and their role in cervical cancer. *Cell Mol Life Sci* **58**(12-13), 1923-42.
- 20) Di Pietro, S. M., Centeno, J. M., Cerutti, M. L., Lodeiro, M. F., Ferreiro, D. U., Alonso, L. G., Schwarz, F. P., Goldbaum, F. A. & De Prat-Gay, G. (2003). Specific Antibody-DNA Interaction: A Novel Strategy for Tight DNA Recognition. *Biochemistry* **42**(20), 6218-27.
- 21) Dill, K. A. (1997). Additivity principles in biochemistry. *J Biol Chem* **272**(2), 701-4.
- 22) Dunitz, J. D. (1995). Win some, lose some: enthalpy-entropy compensation in weak intermolecular interactions. *Chem Biol* **2**(11), 709-12.
- 23) Dyson, H. J. & Wright, P. E. (2002). Coupling of folding and binding for unstructured proteins. *Curr Opin Struct Biol* **12**(1), 54-60.

- 24) Engler, L. E., Sapienza, P., Dörner, L. F., Kucera, R., Schildkraut, I. & Jen-Jacobson, L. (2001). The energetics of the interaction of BamHI endonuclease with its recognition site GGATCC. *J Mol Biol* 307(2), 619-36.
- 25) Fersht, A. (1999). *Structure and mechanism in protein science*, W. H. Freeman & Company, New York.
- 26) Festenstein, R. & Aragon, L. (2003). Decoding the epigenetic effects of chromatin. *Genome Biol* 4(10), 342.
- 27) Foguel, D., Silva, J. L. & de Prat-Gay, G. (1998). Characterization of a partially folded monomer of the DNA-binding domain of human papillomavirus E2 protein obtained at high pressure. *Journal of Biological Chemistry* 273(15), 9050-7.
- 28) Frauenfelder, H. & Leeson, D. T. (1998). The energy landscape in non-biological and biological molecules. *Nat Struct Biol* 5(9), 757-9.
- 29) Frauenfelder, H., Sligar, S. G. & Wolynes, P. G. (1991). The energy landscapes and motions of proteins. *Science* 254(5038), 1598-603.
- 30) Gray, D. (1996). *Circular Dichroism and the Conformational Analysis of Biomolecules* (Fasman, G., Ed.), Plenum Press, New York.
- 31) Guex, N. & Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modelling. *Electrophoresis* 18(15), 2714-23.
- 32) H.M. Berman, J. W., Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne. (2000). The Protein Data Bank. *Nucleic Acids Research* 26, 235-242.
- 33) Ham, J., Dostatni, N., Gauthier, J. M. & Yaniv, M. (1991). The papillomavirus E2 protein: a factor with many talents. *Trends Biochem Sci* 16(11), 440-4.
- 34) Hegde, R., Grossman, S., Lalmins, L. & Sigler, P. (1992). Crystal structure at 1.7Å of the bovine papillomavirus-1 E2 DNA-binding domain bound to its DNA target. *Nature* 359, 505-512.
- 35) Hegde, R. S. (2002). The papillomavirus E2 proteins: structure, function, and biology. *Annu Rev Biophys Biomol Struct* 31, 343-60.

- 36) Hegde, R. S. & Androphy, E. J. (1998). Crystal structure of the E2 DNA-binding domain from human papillomavirus type 16: implications for its DNA binding-site selection mechanism. *Journal of Molecular Biology* **284**(5), 1479-89.
- 37) Hegde, R. S., Wang, A. F., Klm, S. S. & Schapira, M. (1998). Subunit rearrangement accompanies sequence-specific DNA binding by the bovine papillomavirus-1 E2 protein. *Journal of Molecular Biology* **276**(4), 797-808.
- 38) Hill, J. J. & Royer, C. A. (1997). Fluorescence approaches to study of protein-nucleic acid complexation. *Methods In Enzymology* **278**, 390-416.
- 39) Hines, C. S., Meghoo, C., Shetty, S., Biburger, M., Brenowitz, M. & Hegde, R. S. (1998). DNA structure and flexibility in the sequence-specific binding of papillomavirus E2 proteins. *Journal of Molecular Biology* **276**(4), 809-18.
- 40) Hizver, J., Rozenberg, H., Frolov, F., Rabinovich, D. & Shakked, Z. (2001). DNA bending by an adenine—thymine tract and its role in gene regulation. *Proc Natl Acad Sci U S A* **98**(15), 8490-5.
- 41) Hou, S. Y., Wu, S. Y. & Chiang, C. M. (2002). Transcriptional activity among high and low risk human papillomavirus E2 proteins correlates with E2 DNA binding. *J Biol Chem* **277**(47), 45619-29.
- 42) Howley, P. a. L., DR. (2001). *Papillomaviruses and their replication*. Fields Virology (Knipe, D., Howley, PM, Ed.), Lippincott Williams and Wilkins, Philadelphia.
- 43) Itzhaki, L. S., Otzen, D. E. & Fersht, A. R. (1995). The structure of the transition state for folding of chymotrypsin inhibitor 2 analysed by protein engineering methods: evidence for a nucleation-condensation mechanism for protein folding. *J Mol Biol* **254**(2), 260-88.
- 44) Jackson, S. E. & Fersht, A. R. (1991). Folding of chymotrypsin Inhibitor 2. 1. Evidence for a two-state transition. *Biochemistry* **30**(43), 10428-35.
- 45) Jacob, F. a. M., J. (1961). On the regulation of gene activity. *Cold Spring Harbor Symp. Quant. Biol.* **26**, 193-211.

- 46) Jen-Jacobson, L. (1997). Protein-DNA recognition complexes: conservation of structure and binding energy in the transition state. *Biopolymers* 44(2), 153-80.
- 47) Jen-Jacobson, L., Engler, L. E. & Jacobson, L. A. (2000). Structural and thermodynamic strategies for site-specific DNA binding proteins. *Structure Fold Des* 8(10), 1015-23.
- 48) Kim, S. S., Tam, J. K., Wang, A. F. & Hegde, R. S. (2000). The structural basis of DNA target discrimination by papillomavirus E2 proteins. *J Biol Chem* 275(40), 31245-54.
- 49) Kosikov, K. M., Gorin, A. A., Lu, X. J., Olson, W. K. & Manning, G. S. (2002). Bending of DNA by asymmetric charge neutralization: all-atom energy simulations. *J Am Chem Soc* 124(17), 4838-47.
- 50) Kumar, S., Ma, B., Tsal, C. J., Sinha, N. & Nussinov, R. (2000). Folding and binding cascades: dynamic landscapes and population shifts. *Protein Sci* 9(1), 10-9.
- 51) Lai, M. C., Teh, B. H. & Tarn, W. Y. (1999). A human papillomavirus E2 transcriptional activator. The interactions with cellular splicing factors and potential function in pre-mRNA processing. *J Biol Chem* 274(17), 11832-41.
- 52) Lakowicz, J. R. (1999). *Principles of fluorescence spectroscopy*, Kluwer Academic/Plenum Publishers, New York.
- 53) Lewontin, R. C. (2000). *The Dream of the Human Genome and Other Illusions*, New York Review Books, New York.
- 54) Liang, H., Petros, A. M., Meadows, R. P., Yoon, H. S., Egan, D. A., Walter, K., Holzman, T. F., Robins, T. & Fesik, S. W. (1996). Solution structure of the DNA-binding domain of a human papillomavirus E2 protein: evidence for flexible DNA-binding regions. *Biochemistry* 35(7), 2095-103.
- 55) Lima, L. M. & de Prat-Gay, G. (1997). Conformational changes and stabilization induced by ligand binding in the DNA-binding domain of the E2 protein from human papillomavirus. *Journal of Biological Chemistry* 272(31), 19295-303.

- 56) Lima, L. M., Foguel, D. & Silva, J. L. (2000). DNA tightens the dimeric DNA-binding domain of human papillomavirus E2 protein without changes in volume. *Proc Natl Acad Sci U S A* 97(26), 14289-94.
- 57) Lin, S. & Riggs, A. D. (1975). The general affinity of lac repressor for E. coli DNA: implications for gene regulation in procaryotes and eucaryotes. *Cell* 4(2), 107-11.
- 58) Livingstone, J. R., Spolar, R. S. & Record, M. T., Jr. (1991). Contribution to the thermodynamics of protein folding from the reduction in water-accessible nonpolar surface area. *Biochemistry* 30(17), 4237-44.
- 59) Lowy, D., and Howley, PM. (2001). *Papillomaviruses*. Fields Virology (Knipe, D., Howley, PM, Ed.), Lippincott Williams and Wilkins, Philadelphia.
- 60) Lundback, T. & Hard, T. (1996). Sequence-specific DNA-binding dominated by dehydration. *Proc Natl Acad Sci U S A* 93(10), 4754-9.
- 61) Luscombe, N. M., Austin, S. E., Berman, H. M. & Thornton, J. M. (2000). An overview of the structures of protein-DNA complexes. *Genome Biol* 1(1), REVIEWS001.
- 62) Luscombe, N. M., Laskowski, R. A. & Thornton, J. M. (2001). Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level. *Nucleic Acids Res* 29(13), 2860-74.
- 63) Luscombe, N. M. & Thornton, J. M. (2002). Protein-DNA interactions: amino acid conservation and the effects of mutations on binding specificity. *J Mol Biol* 320(5), 991-1009.
- 64) Martin, T. W. & Derewenda, Z. S. (1999). The name is bond, H bond. *Nat Struct Biol* 6(5), 403-6.
- 65) Massimi, P., Pim, D., Bertoli, C., Bouvard, V. & Banks, L. (1999). Interaction between the HPV-16 E2 transcriptional activator and p53. *Oncogene* 18(54), 7748-54.
- 66) Matsumoto, T., Nakashima, N., Takase, K., Hirochika, H. & Mizuno, H. (1997). A mutation study of the DNA binding domain of human papillomavirus type11 E2 protein. *J Biochem (Tokyo)* 121(1), 138-44.

- 67) Matthews, B. W. (1988). Protein-DNA interaction. No code for recognition. *Nature* **335**(6188), 294-5.
- 68) McBride, A., Romanczuk, H. & Howley, P. (1991). The papillomavirus E2 regulatory proteins. *Journal of Biological Chemistry* **266**(28), 18411-18414.
- 69) McBride, A. A., Byrne, J. C. & Howley, P. M. (1989). E2 polypeptides encoded by bovine papillomavirus type 1 form dimers through the common carboxyl-terminal domain: transactivation is mediated by the conserved amino-terminal domain. *Proc Natl Acad Sci U S A* **86**(2), 510-4.
- 70) McKay, D. B. & Steitz, T. A. (1981). Structure of catabolite gene activator protein at 2.9 Å resolution suggests binding to left-handed B-DNA. *Nature* **290**(5809), 744-9.
- 71) Mok, Y. K., Alonso, L. G., Lima, L. M., Bycroft, M. & de Prat-Gay, G. (2000). Folding of a dimeric beta-barrel: residual structure in the urea denatured state of the human papillomavirus E2 DNA binding domain. *Protein Sci* **9**(4), 799-811.
- 72) Mok, Y. K., Bycroft, M. & de Prat-Gay, G. (1996). The dimeric DNA binding domain of the human papillomavirus E2 protein folds through a monomeric intermediate which cannot be native-like. *Nature Structural Biology* **3**(8), 711-7.
- 73) Mok, Y. K., de Prat Gay, G., Butler, P. J. & Bycroft, M. (1996). Equilibrium dissociation and unfolding of the dimeric human papillomavirus strain-16 E2 DNA-binding domain. *Protein Science* **5**(2), 310-9.
- 74) Muller, F. & Sapp, M. (1996). Domains of the E1 protein of human papillomavirus type 33 involved in binding to the E2 protein. *Virology* **219**(1), 247-56.
- 75) Oda, M. & Nakamura, H. (2000). Thermodynamic and kinetic analyses for understanding sequence-specific DNA recognition. *Genes Cells* **5**(5), 319-26.
- 76) Ohlendorf, D. H., Anderson, W. F., Fisher, R. G., Takeda, Y. & Matthews, B. W. (1982). The molecular basis of DNA-protein recognition inferred from the structure of cro repressor. *Nature* **298**(5876), 718-23.
- 77) Oobatake, M., Kono, H., Wang, Y. & Sarai, A. (2003). Anatomy of specific interactions between lambda repressor and operator DNA. *Proteins* **53**(1), 33-43.

78) Ozkan, S. B., Bahar, I. & Dill, K. A. (2001). Transition states and the meaning of Phi-values in protein folding kinetics. *Nat Struct Biol* 8(9), 765-9.

79) Pabo, C. O. & Lewis, M. (1982). The operator-binding domain of lambda repressor: structure and DNA recognition. *Nature* 298(5873), 443-7.

80) Pabo, C. O. & Nekludova, L. (2000). Geometric analysis and comparison of protein-DNA interfaces: why is there no simple code for recognition? *J Mol Biol* 301(3), 597-624.

81) Pace, C. N., Vajdos, F., Lanette, F., Grimsley, G. & Gray, T. (1995). How to measure and predict the molar absorption coefficient of a protein. *Protein Science* 4, 2411-23.

82) Phelps, W. C., Barnes, J. A. & Lobe, D. C. (1998). Molecular targets for human papillomaviruses: prospects for antiviral therapy. *Antivir Chem Chemother* 9(5), 359-77.

83) Powell, R. M., Parkhurst, K. M., Brenowitz, M. & Parkhurst, L. J. (2001). Marked stepwise differences within a common kinetic mechanism characterize TATA-binding protein interactions with two consensus promoters. *J Biol Chem* 276(32), 29782-91.

84) Powell, R. M., Parkhurst, K. M. & Parkhurst, L. J. (2002). Comparison of TATA-binding protein recognition of a variant and consensus DNA promoters. *J Biol Chem* 277(10), 7776-84.

85) Record, M. T., Jr., Ha, J. H. & Fisher, M. A. (1991). Analysis of equilibrium and kinetic measurements to determine thermodynamic origins of stability and specificity and mechanism of formation of site-specific complexes between proteins and helical DNA. *Methods Enzymol* 208, 291-343.

86) Rodnina, M. V. & Wintermeyer, W. (2001). Ribosome fidelity: tRNA discrimination, proof-reading and induced fit. *Trends Biochem Sci* 26(2), 124-30.

87) Sanders, C. M. & Maitland, N. J. (1994). Kinetic and equilibrium binding studies of the human papillomavirus type-16 transcription regulatory protein E2 interacting with core enhancer elements. *Nucleic Acid Research* 22(22), 4890-97.

88) Sanders, C. M. & Stenlund, A. (2001). Mechanism and requirements for bovine papillomavirus, type 1, E1 initiator complex assembly promoted by the E2 transcription factor bound to distal sites. *J Biol Chem* 276(26), 23689-99.

- 89) Schreiber, G. (2002). Kinetic studies of protein-protein interactions. *Curr Opin Struct Biol* 12(1), 41-7.
- 90) Seedorf, K., Krammer, G., Durst, M., Suhai, S. & Rowekamp, W. G. (1985). Human papillomavirus type 16 DNA sequence. *Virology* 145(1), 181-5.
- 91) Shoemaker, B. A., Portman, J. J. & Wolynes, P. G. (2000). Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. *Proc Natl Acad Sci U S A* 97(16), 8868-73.
- 92) Shope, R. E. a. H., E.W. (1933). Infectious papillomatosis of rabbits; with a note on the histopathology. *J. Exp. Med.* 58, 607-624.
- 93) Spalholz, B. A., Lambert, P. F., Yee, C. L. & Howley, P. M. (1987). Bovine papillomavirus transcriptional regulation: localization of the E2-responsive elements of the long control region. *J Virol* 61(7), 2128-37.
- 94) Spalholz, B. A., Yang, Y. C. & Howley, P. M. (1985). Transactivation of a bovine papilloma virus transcriptional regulatory element by the E2 gene product. *Cell* 42(1), 183-91.
- 95) Spolar, R. S., Livingstone, J. R. & Record, M. T., Jr. (1992). Use of liquid hydrocarbon and amide transfer data to estimate contributions to thermodynamic functions of protein folding from the removal of nonpolar and polar surface from water. *Biochemistry* 31(16), 3947-55.
- 96) Spolar, R. S. & Record, M. T., Jr. (1994). Coupling of local folding to site-specific binding of proteins to DNA. *Science* 263(5148), 777-84.
- 97) Strogatz, S. H. (2001). Exploring complex networks. *Nature* 410(6825), 268-76.
- 98) Stryer, L. (1995). *Biochemistry*. 4th edit, W.H. Freeman and Co, New York.
- 99) Takeda, Y., Sarai, A. & Rivera, V. M. (1989). Analysis of the sequence-specific interactions between Cro repressor and operator DNA by systematic base substitution experiments. *Proc Natl Acad Sci U S A* 86(2), 439-43.
- 100) Thain, A., Webster, K., Emery, D., Clarke, A. R. & Gaston, K. (1997). DNA binding and bending by the human papillomavirus type 16 E2 protein. Recognition of an extended binding site. *Journal of Biological Chemistry* 272(13), 8236-42.

- 101) Tsai, C. J., Ma, B. & Nussinov, R. (1999). Folding and binding cascades: shifts in energy landscapes. *Proc Natl Acad Sci U S A* 96(18), 9970-2.
- 102) Tsodikov, O. V., Record, M. T., Jr. & Sergeev, Y. V. (2002). Novel computer program for fast exact calculation of accessible and molecular surface areas and average surface curvature. *J Comput Chem* (6), 600-9.
- 103) Turazza, E., Lapena, A., Sprovieri, O., Torres, C. P., Gurucharri, C., Maciel, A., Lema, B., Grinstein, S. & Kahn, T. (1997). Low-risk human papillomavirus types 6 and 11 associated with carcinomas of the genital and upper aero-digestive tract. *Acta Obstet Gynecol Scand* 76(3), 271-6.
- 104) Veeraraghavan, S., Mello, C. C., Androphy, E. J. & Baleja, J. D. (1999). Structural correlates for enhanced stability in the E2 DNA-binding domain from bovine papillomavirus. *Biochemistry* 38(49), 16115-24.
- 105) von Hippel, P. & Berg, O. (1986). On the specificity of DNA-protein interactions. *Proceedings of the National Academy of Sciences of the United States of America* 83(6), 1608-1612.
- 106) von Hippel, P. H. & Berg, O. G. (1989). Facilitated target location in biological systems. *J Biol Chem* 264(2), 675-8.
- 107) Watson, J. D., Gilman M., Witkowski J., Zoller, M. (1992). *Recombinant DNA*, W.H. Freeman, New York.
- 108) Watson, J. D. C., F.H.C. (1953). A structure for desoxyribose nucleic acid. *Nature* 171(737-738).
- 109) Wells, J. A. (1991). Systematic mutational analyses of protein-protein interfaces. *Methods Enzymol* 202, 390-411.
- 110) Wilson, V. G., West, M., Woytek, K. & Rangasamy, D. (2002). Papillomavirus E1 proteins: form, function, and features. *Virus Genes* 24(3), 275-90.
- 111) Winter, R. B., Berg, O. G. & von Hippel, P. H. (1981). Diffusion-driven mechanisms of protein translocation on nucleic acids. 3. The Escherichia coli lac repressor—operator interaction: kinetic measurements and conclusions. *Biochemistry* 20(24), 6961-77.
- 112) Zimmerman, J. M. & Maher, L. J., 3rd. (2003). Solution measurement of DNA curvature in papillomavirus E2 binding sites. *Nucleic Acids Res* 31(17), 5134-9.