

UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS

INSTITUTO DE INVESTIGACIÓN EN CIENCIAS BÁSICAS Y APLICADAS

CENTRO DE INVESTIGACIÓN EN DINÁMICA CELULAR

Dinámica molecular de la recombinasa Cre asociada al reconocimiento de *loxP*

TESIS PROFESIONAL PARA OBTENER EL GRADO DE:

DOCTORADO EN CIENCIAS

PRESENTA:

MARCO ANTONIO

RAMÍREZ MARTÍNEZ

DIRECTOR DE TESIS: DRA. CARMEN NINA PASTOR COLÓN

CUERNAVACA, MORELOS MAYO, 2025

Instituto de Investigación en Ciencias Básicas y Aplicadas

UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS

INSTITUTO DE INVESTIGACIÓN EN CIENCIAS BÁSICAS Y APLICADAS

CENTRO DE INVESTIGACIÓN EN DINÁMICA CELULAR

Dinámica molecular de la recombinasa Cre asociada al reconocimiento de *loxP*

TESIS PROFESIONAL PARA OBTENER EL GRADO DE:

DOCTORADO EN CIENCIAS

PRESENTA:

MARCO ANTONIO

RAMÍREZ MARTÍNEZ

DIRECTOR DE TESIS: DRA. CARMEN NINA PASTOR COLÓN

CUERNAVACA, MORELOS MAYO, 2025

Instituto de Investigación en Ciencias Básicas y Aplicadas

Agradecimientos

Al Laboratorio de Supercómputo y Visualización en Paralelo de la UAM Iztapalapa a través de los proyectos 2021 y 2023; al Laboratorio Nacional del Supercómputo por los recursos otorgados con el proyecto 202301027N. A la Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) por la beca otorgada (CVU:858905), así como por los proyectos INFR-2014-02-231504, A1-S-11842, y CBF2023-2024-3670.

Índice

Lista de Figuras ***5***

Lista de Tablas ***7***

Abreviaturas ***7***

Resumen ***8***

Abstract ***10***

1. Introducción ***12***

1.1 Tirosina recombinasas (Y-SSR) **12**

1.2 Sistema Cre-*loxP* **15**

1.2.1 Bacteriófago P1 15

1.2.2 Recombinasa Cre 16

1.2.3 Secuencia *loxP* y armado del intasoma 17

1.4 Métodos computacionales **21**

1.4.1 Generalidades de la dinámica molecular 22

2. Antecedentes ***23***

2.1.1 Estudios estructurales del complejo Cre-*loxP* 23

2.1.2 Dinámica molecular del sistema Cre-*loxP* 29

3. Justificación ***30***

4. Hipótesis ***30***

5. Objetivo ***31***

5.1 Objetivos particulares **31**

6. Metodología *31*

6.1 Obtención de estructuras iniciales **31**

6.2 Optimización de estructuras **33**

6.3 Construcción de celdas de simulación **35**

6.4 Dinámica molecular **37**

6.5 Cálculo de propiedades estructurales **40**

6.5.1 Radio de giro y RMSD 41

6.5.2 Ángulos entre hélices 41

6.5.3 Contactos entre átomos pesados 41

6.5.4 Estructura secundaria 42

7. Resultados *43*

7.1 Dímero 7RHZ **43**

7.1.1 Armado del complejo y validación experimental 43

7.1.2 Formación y ruptura de interfaces 45

7.1.3 Dinámica de *loxI* y la posición de la hélice N 48

7.1.4 Interruptores moleculares 54

7.1.5 Integridad del sitio activo 57

7.2 Cre libre y dominios aislados **61**

7.2.1 Validación experimental 61

7.2.2 Potencial electrostático y radio de giro 62

7.2.3 Flexibilidad del conector entre dominios 63

7.2.4 Múltiples orientaciones entre dominios 65

7.2.5 Accesibilidad de superficies de interacción 67

7.2.6 Carácter desordenado del extremo C-terminal 69

7.2.7 El conector y el extremo C-terminal bloquean al sitio activo 71

7.2.8 Integridad del sitio activo en ausencia de *loxI* 72

7.2.9 La hélice A muestrea múltiples estados 78

7.2.10 “Bisagra” del dominio CAT 83

7.2.11 Estructura de CBD aislado y CBD-*loxA* 84

8. Conclusiones 88

9. Perspectivas 91

10. Bibliografía 94

11. Anexos 100

Lista de Figuras

Figura 1. Mecanismo de reacción de las Y-SSR……13

Figura 2. Sistema Cre-*loxP*……16

Figura 3. Estructura de Cre y dominio CAT ……..17

Figura 4. Secuencia *loxP*……..……..18

Figura 5. Vía de ensamble del intasoma…….19

Figura 6. Interacciones Cre-*loxP* y mutantes de Cre..……..20

Figura 7. Interfaces proteína-proteína…….……21

Figura 8. "Sustrato suicida" para usado para capturar el intermediario covalente de Cre…….....24

Figura 9. Diferentes posiciones de la hélice A…….29

Figura 10. Secuencias *lox* y restricciones……37

Figura 11. Estructura del dímero 7RHZ completo…..……44

Figura 12. Comparación del patrón de protección experimental con los contactos proteína-azúcar…….45

Figura 13. Formación y rompimiento de interfaces……...46

Figura 14. Invasión de la hélice N hacia el monómero adyacente..……48

Figura 15. Representación esquemática del parámetro *roll*…….49

Figura 16. Valor de *roll* asociado a la posición de la hélice N..……...51

Figura 17. Valor de *roll* de *loxI* independiente a la hélice N……..……..52

Figura 18. Doblez de *loxI* por cambio en el valor de *roll*…..……..53

Figura 19. Desplazamiento de la hélice A……..55

Figura 20. Paisaje conformacional de las hélices I-K……..……..56

Figura 21. Ensamble del sitio activo…….……59

Figura 22. Detalle de las interacciones con el sitio activo …..…….60

Figura 23. Comparación con datos experimentales: desplazamientos químicos de Cα y Cβ ……..……...62

Figura 24. Potencial electrostático y distribución de radio de giro……..63

Figura 25. Mapa de Ramachandran y distribución de distancia entre extremos del conector……. ……64

Figura 26. Paisaje conformacional de Cre libre y con *loxI*…..…….66

Figura 27. Superficies de interacción proteína-proteína y proteína-ADN…..……..67

Figura 28. Exposición de áreas de interacción…..……..69

Figura 29. Interacciones del extremo C-terminal y mapa de Ramachandran……..71

Figura 30. Múltiples interacciones del extremo C-terminal y el conector…..……..72

Figura 31. Distribución de radio de giro del sitio activo……..……..74

Figura 32. Distribución de ángulos diedros de cadena lateral……..78

Figura 33. Definición de ángulos χ1 y χ2 ……..……..78

Figura 34. Paisaje conformacional de la hélice A en diferentes contextos..……..80

Figura 35. Desplazamiento de la hélice A ……..81

Figura 36. Transiciones entre estados …….……...82

Figura 37. Paisaje conformaciones de las hélices I-K para todos los sistemas con dominio CAT ……..……...84

Figura 38. Espectro de dicroísmo circular ……..85

Figura 39. Estructura secundaria de la hélice A con y sin ADN ……..87

Lista de Tablas

Tabla 1. Secuencias blanco de Y-SSR ……14

Tabla 2. Estructuras de Cre disponibles en el PDB ……24

Tabla 3. Colección de estructuras iniciales ……..32

Tabla 4. Tamaño de los sistemas obtenidos en el servidor charmm-gui ……..35

Tabla 5. Tiempo acumulado de simulación para cada sistema ……39

Abreviaturas

MD = Dinámica Molecular (por sus siglas en inglés)

IDR = Región intrínsecamente desordenadas (por sus siglas en inglés)

1Q3U:A = Cadena A (monómero inactivo) de la estructura con PDBID 1Q3U

1Q3U:F = Cadena F (monómero activo) de la estructura con PDBID 1Q3U

DC = Dicroísmo circular

CS = Desplazamientos químicos (por sus siglas en inglés)

μs = microsegundo ($1X10^{-6}$ segundos)

ns = nanosegundo ($1X10^{-9}$ segundos)

ps = picosegundos ($1X10^{-12}$ segundos)

fs = femtosegundos ($1X10^{-15}$ segundos)

RMN = Resonancia Magnética Nuclear

ADN = Ácido desoxirribonucleico

Resumen

La recombinasa Cre forma un complejo tetramérico (intasoma) de monómeros activos e inactivos alternados, que se unen a dos secuencias de ADN conocidas como *loxP.* Este complejo se estabiliza mediante interacciones proteína-proteína y proteína-ADN. Este sistema, proveniente del bacteriófago P1, realiza dos actividades principales en el contexto de la infección viral: 1) circularización del genoma viral, lo que permite a P1 coexistir como un plásmido dentro de la bacteria que infecta, y 2) la resolución de dímeros de ADN viral formados durante la replicación del genoma viral durante división celular. Cada monómero de Cre está compuesto por dos dominios (CBD y CAT). Cre no necesita cofactores o proteínas accesorias para llevar a cabo la reacción de recombinación, lo que ha convertido al sistema Cre-*loxP* en una herramienta poderosa para la ingeniería genética. Sin embargo, aún falta información estructural sobre Cre en ausencia de ADN.

En este estudio, realizamos simulaciones de dinámica molecular (MD, por sus siglas en inglés) de cinco estructuras de Cre, provenientes de diferentes etapas del ensamblaje del intasoma. Se realizaron cuatro réplicas de 5 μs cada una, lo que dio lugar a un tiempo total de simulación de 100 μs, siendo esta la simulación de MD más larga realizada para este sistema. Además, se evaluó el efecto de la presencia de ADN en la dinámica de Cre utilizando simulaciones de un dímero de Cre unido a un sitio *loxI*, el dominio CBD en presencia/ausencia ADN y el dominio CAT en presencia/ausencia de la hélice N. La dinámica de Cre libre muestra que el conector entre dominios es una región intrínsecamente desordenada (IDR, por sus siglas en inglés) que favorece diferentes orientaciones de los dos dominios. A pesar del aumento en la flexibilidad, los dominios permanecen plegados y hay pocos contactos entre ellos, esto correlaciona con datos experimentales. La hélice N en el dominio CAT también es una IDR que interactúa con toda la proteína, específicamente con el sitio activo, sugiriendo la formación de un complejo auto-inhibido. El dominio CAT exhibe un movimiento de apertura y cierre, similar al de una almeja, lo que podría ser relevante para establecer interacciones proteína-

proteína en el contexto de un complejo dímero y/o intasoma. Por otro lado, la hélice A en el dominio CBD muestra un movimiento similar al de un limpiaparabrisas, explorando conformaciones tipo monómero/dímero e intasoma. Al comparar este comportamiento con los dominios aislados y con un dímero de Cre unido a un sitio *loxI*, se observó que los movimientos observados en la proteína completa, así como el plegamiento de ambos dominios, se conservan. Finalmente, esta flexibilidad se reduce drásticamente en el dímero unido a *loxI*, lo que indica que la dinámica de Cre depende del contexto en el que se encuentre. Este estudio arroja luz sobre la dinámica de Cre en ausencia de ADN y proporciona una explicación para la ausencia de una estructura experimental de Cre en este estado.

El acceso a la publicación asociada a esta tesis se muestra en el Anexo I

Abstract

Cre recombinase forms a tetrameric complex (intasome) consisting of alternating active and inactive monomers, which bind to two DNA sequences known as *loxP*. This complex is stabilized by protein–protein and protein–DNA interactions. Originating from bacteriophage P1, this system performs two main functions during viral infection: (1) circularization of the viral genome, allowing P1 to persist as a plasmid within the host bacterium, and (2) resolution of viral DNA dimers formed during genome replication in cell division. Each Cre monomer is composed of two domains (CBD and CAT). Cre does not require cofactors or accessory proteins to carry out the recombination reaction, which has made the Cre-*loxP* system a powerful tool in genetic engineering. However, structural information about Cre in the absence of DNA is still lacking.

In this study, we conducted molecular dynamics (MD) simulations of five Cre structures corresponding to different stages of intasome assembly. Four replicas of 5 μs each were performed, resulting in a total simulation time of 100 μs, making this the longest MD simulation conducted for this system to date. Additionally, the effect of DNA presence on Cre dynamics was evaluated using simulations of a Cre dimer bound to a *loxI* site, the CBD domain with and without DNA, and the CAT domain with and without the helix N. Simulations of free Cre revealed that the linker between domains is an intrinsically disordered region (IDR) that promotes various orientations of the two domains. Despite increased flexibility, the domains remain folded and show few inter-domain contacts, which is consistent with experimental data. The N-helix within the CAT domain is also an IDR that interacts with the entire protein, particularly with the active site, suggesting the formation of an autoinhibited complex. The CAT domain exhibits a clam-like opening and closing motion, potentially relevant for establishing protein–protein interactions in the context of a dimer or intasome complex. Meanwhile, the helix A in the CBD domain shows a windshield-wiper-like motion, sampling monomeric/dimeric and intasome-like conformations. Comparison with the isolated domains and a Cre dimer bound to a

loxI site showed that the motions observed in the full-length protein, as well as the folding of both domains, are preserved. Finally, this flexibility is drastically reduced in the loxI-bound dimer, indicating that Cre dynamics are context-dependent. This study sheds light on the dynamics of Cre in the absence of DNA and provides an explanation for the lack of an experimental structure of Cre in this state.

The access to the publication associated with this thesis is shown in Appendix I

1. Introducción

1.1 Tirosina recombinasas (Y-SSR)

Las tirosina recombinasas sitio específicas (Y-SSR) son proteínas que se unen al ADN, y promueven su corte, recombinación y ligación en secuencias específicas[1]. El nombre proviene del residuo que usan para llevar a cabo la reacción de recombinación[2].

Para llevar a cabo la reacción, se debe formar un tetrámero de monómeros activos e inactivos alternados, que se unen a sus secuencias específicas[3]. El mecanismo de acción está compartido entre Y-SSRs, éste se divide en seis pasos, los cuales se describen a continuación y se muestran esquemáticamente en la Figura 1[4]:

1. **Sinapsis**: cuatro Y-SSRs interactúan con dos sustratos de ADN para formar un complejo tetramérico (intasoma). El complejo está formado por dos monómeros en conformación activa (listos para llevar a cabo el ataque nucleofílico); y dos en conformación inactiva.
2. **Corte/ligación**: la tirosina catalítica realiza un ataque nucleofílico sobre el fosfato escindible de la región de recombinación presente en el ADN, formando un enlace 3'-fosfotirosina y dejando un grupo 5'-OH libre.
3. **Intercambio de cadena**: el 5'-OH libre ataca al enlace 3'-fosfotirosina del otro dúplex de DNA llevando a la formación de una estructura de Holliday.
4. **Isomerización**: los monómeros activos cambian a la conformación inactiva; mientras que, los inactivos pasan al estado activo, dejando al complejo listo para otra ronda de corte e intercambio.
5. **Segundo ataque nucleofílico:** la tirosina catalítica ataca al nuevo fosfato escindible en el intermediario de Holliday, generando un nuevo enlace de fosfotirosina y dejando nuevamente un extremo 5'-OH libre.
6. **Resolución de la estructura de Holliday:** el extremo 5'-OH libre ataca al enlace fosfotirosina llevando a la obtención del blanco recombinado.

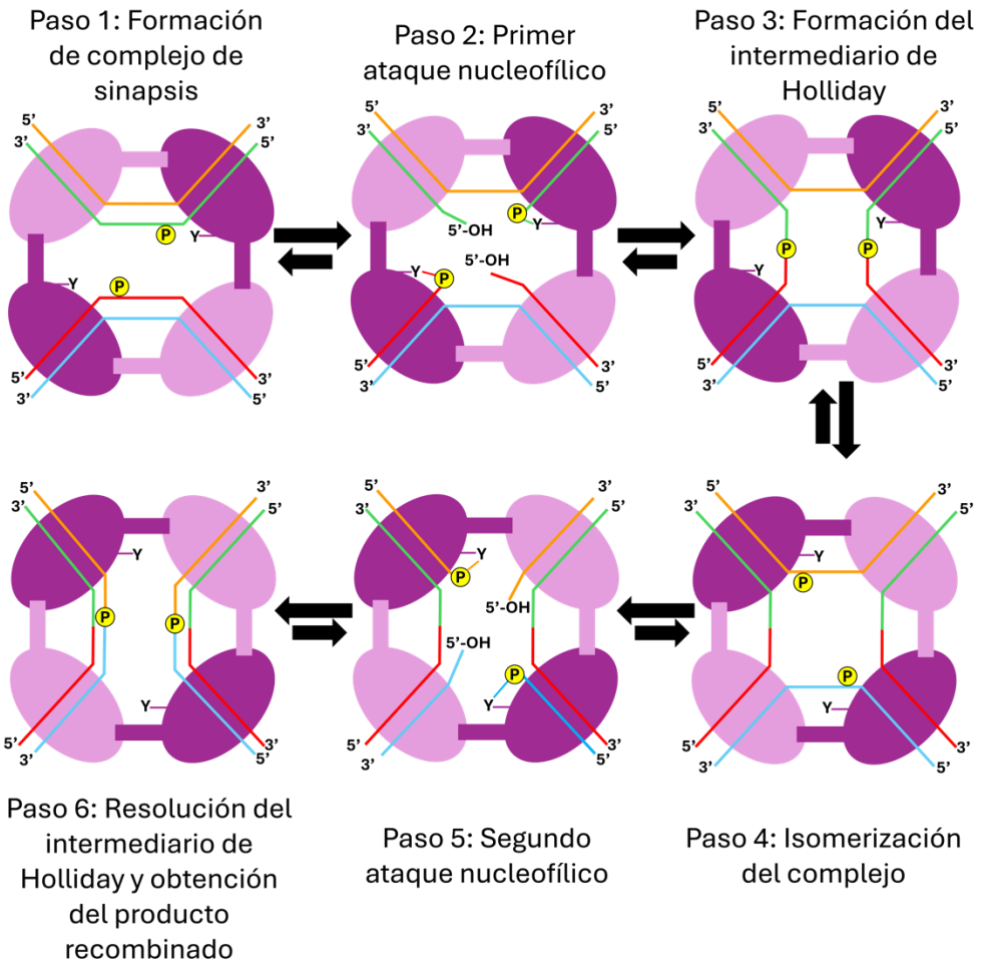


Figura 1. Mecanismo de reacción de las Y-SSR. Los monómeros activos se presentan en morado; mientras que, los inactivos están en rosa. La tirosina catalítica se muestra en negritas; mientras que, el fosfato escindible está encerrado en amarillo. Cada una de las hebras de ADN se muestra en diferente color.

Cada recombinasa se une a una secuencia blanco específica. Estas secuencias varían en longitud, pero están compuestas por dos regiones de ADN que flanquean a una región espaciadora, la cual es la zona en la que el corte y empalme suceden. Esto ha permitido el uso combinado de Y-SSR para la edición de genomas sin que haya reconocimiento cruzado entre recombinasas[2]. En la tabla 1 se muestran las secuencias que reconocen tres de las Y-SSR más estudiadas hasta el momento.

Cada una de estas recombinasas proviene de distintos organismos, pero pertenecen a la misma familia[2] (Flp: *Saccaromyces cerevisiae*, Cre: bacteriófago P1, Int: bacteriófago lambda).

Tabla 1. Secuencias blanco de Y-SSR. Para cada Y-SSR se muestra la secuencia que reconoce. En rojo, se marca la región de corte.

Sistema	Secuencia blanco
Flp/*FRT*	5'-GAACTTCCTATACTTTCTAGAGAATAGGAACTTC-3'
Cre/*loxP*	5'-ATAACTTCGTATAATGTATGCTATACGAAGTTAT-3'
Int/*attP*	5'-CAGTTTTTTATACTAAGTTG-3'

Gracias a su facilidad de uso y al hecho de que no necesita cofactores o el uso de proteínas accesorias[5] (como el sistema Int/*attP*), el sistema Cre-*loxP* ha sido ampliamente utilizado en aplicaciones biotecnológicas[6]. Además, a diferencia de Flp/*FRT*, el cual necesita de la formación del complejo tetramérico para poder armar un sitio activo completo[2], Cre tiene un sitio activo completo en cada monómero.

A continuación, se describen sus características principales, tanto estructurales como funcionales, así como las preguntas que siguen sin ser del todo entendidas y que inspiran este trabajo.

1.2 Sistema Cre-*loxP*

1.2.1 Bacteriófago P1

El bacteriófago P1, Infecta a *Escherichia coli* (*E. coli*) y otras bacterias entéricas. Su estructura (similar a la del bacteriófago T4) consta de una cabeza icosaédrica, unida a un tallo contráctil del que salen seis fibras, estas sirven de anclaje a la bacteria[7]. A diferencia de otros bacteriófagos que se insertan en el cromosoma del huésped (fago lambda) durante su fase lisogénica, P1 existe como un plásmido independiente dentro de *E. coli.* Este mecanismo requiere de la circularización del genoma viral, proceso que es llevado a cabo por el sistema Cre-*loxP*[8] (Figura 2A) Se ha estimado que aproximadamente 1 de cada 10^4 células bacterianas pierden el plásmido viral debido a la formación de dímeros de ADN generados durante la replicación que no fueron resueltos por la maquinaria celular[9,10]. El sistema Cre-*loxP* también participa en este proceso promoviendo el corte y repartición equitativa del plásmido viral a las bacterias hijas (Figura 2B).

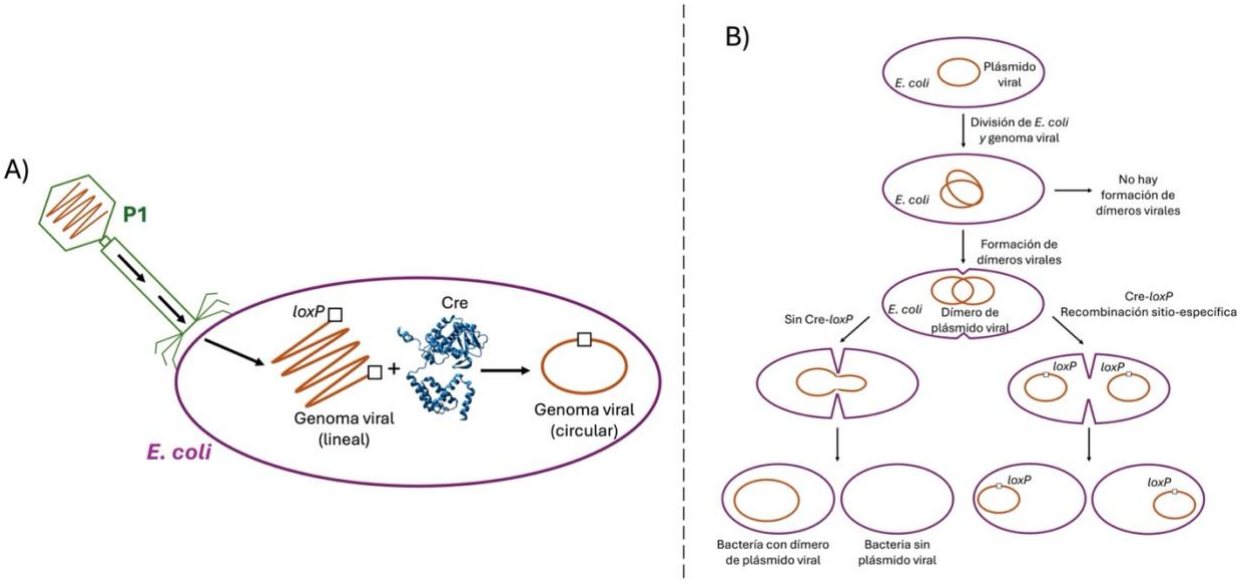


Figura 2. Sistema Cre-*loxP.* A) Representación esquemática de P1 insertando su genoma en *E. coli* y subsecuente circularización de este por Cre. B) Separación de

dímeros de plásmido viral mediada por Cre. Para mejor visualización, el genoma de *E. coli* no se muestra en la figura.

1.2.2 Recombinasa Cre

La recombinasa Cre (*Causes recombination*) es una proteína de 343 residuos que se expresa en las primeras etapas de la infección viral. Esta proteína está formada por dos dominios (*Core Binding Domain*, CBD; y *Catalytic Domain*, CAT, por sus siglas en inglés) unidos entre sí por un conector flexible[3]. El dominio CBD comprende las hélices A a la E; mientras que, el dominio CAT comprende las hélices F a la N y dos hojas beta pequeñas (Figura 3). Además, este dominio contiene al residuo catalítico (Y324), y al resto de los aminoácidos del sitio activo[3].

Cada monómero de Cre se une a *loxP* e interactúa con ambos surcos del ADN, formando una estructura tipo "tenaza". Esto resulta en una amplia superficie de interaccion entre Cre y el *loxP*[2]. Como se mencionó anteriormente, es necesaria la formación de un intasoma, el cual, es estabilizado por múltiples interacciones. Cre existe en dos conformaciones, activa e inactiva; estas conformaciones intercambian entre sí a lo largo de la reacción (Figura 1). Las principales diferencias entre ambas conformaciones ocurren en el dominio CAT, en donde la horquilla β2/3 pasa de tener una estructura tipo beta plegada a estructuras irregulares; esta horquilla es de interés ya que en ella reside la Lisina 201, el cual participa en la reacción de recombinación[11,12]. En el caso del dominio CBD, no se han descrito cambios estructurales asociados al cambio de conformación, y se cree que este dominio es

necesario para la correcta unión de Cre a *loxP* durante los primeros pasos del ensamblaje del complejo[6].

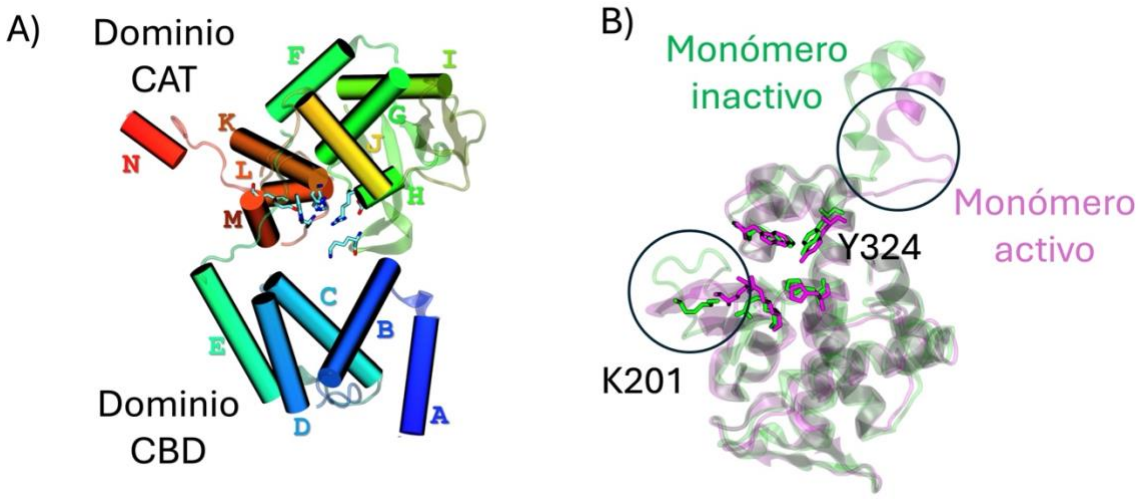


Figura 3. Estructura de Cre y dominio CAT. A) Monómero de Cre obtenido de la estructura con PDBID 1Q3U[13]. Cada hélice está etiquetada con el color que corresponde a su color en la estructura en la escala BGR. B) Sobrelape de la conformación activa (PDB-ID 1Q3U:F, magenta) e inactiva (PDB-ID 1Q3U:A, verde) del dominio CAT; los círculos muestran las zonas con más diferencias. Las cadenas laterales de los cuatro residuos del sitio activo que más cambian también se muestran como varillas.

1.2.3 Secuencia *loxP* y armado del intasoma

loxP es una secuencia de 34 pares de bases (pb) que consta de dos secuencias de 13 pb cada una denominadas Elementos de Unión de la Recombinasa (RBE, por sus siglas en inglés), y un espaciador de 8 pares de bases localizado en medio de los dos RBE[8] (Figura 4). Es en esta región en donde el corte y ligación ocurren, además, la orientación de la secuencia espaciadora es la que le da direccionalidad a la reacción[14].

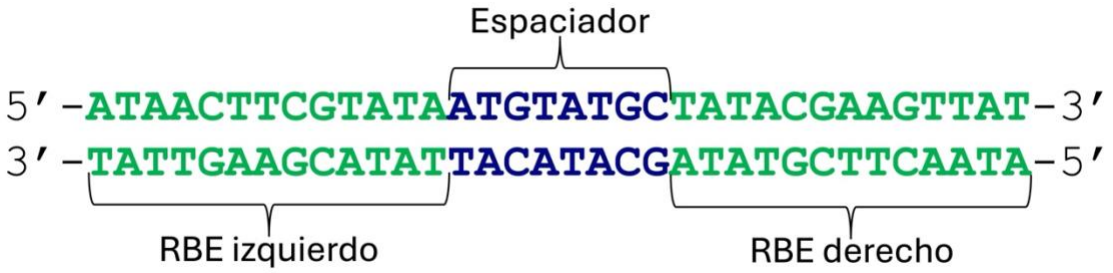


Figura 4. Secuencia *loxP*. En verde se muestra la secuencia de cada RBE mientras que en azul está marcado el espaciador.

La unión de Cre a *loxP* es un proceso cooperativo, en el cual un monómero de Cre se une a cada RBE, el primero con afinidad nanomolar y el segundo con afinidad picomolar[14]. Esto lleva a la formación de un complejo Cre_2-*loxP*. Posteriormente, dos de estos complejos se unen entre sí (afinidad nanomolar) para formar el intasoma (complejo Cre_4-$loxP_2$), la cual es la unidad funcional de la recombinación[6,14,15] (Figura 5).

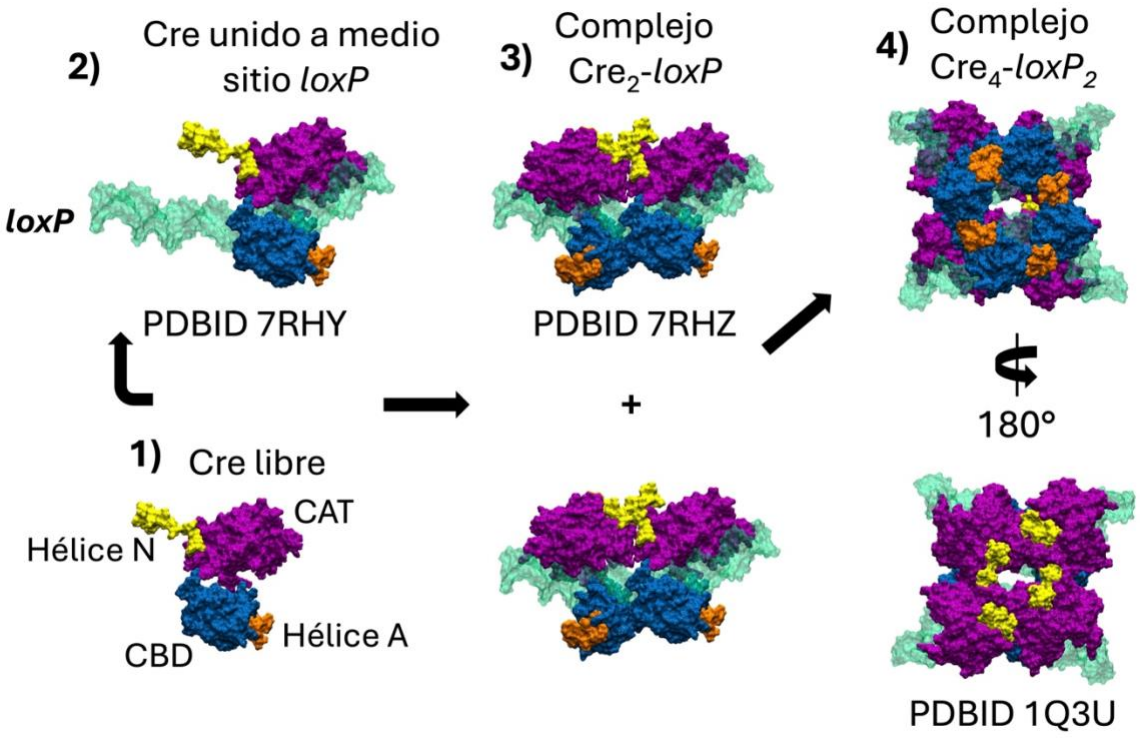


Figura 5. Vía de ensamble del intasoma. 1) Cre libre (sin estructura experimental disponible), los dominios CBD Y CAT se muestran en azul y morado, respectivamente; las hélices A (anaranjado) y N (amarillo) también se muestran en la estructura. 2) Monómero de Cre unido a medio sitio *loxP* (verde) (PDBID 7RHY[14]), un segundo monómero se une al RBE libre llevando a la formación del 3) complejo Cre_2-*loxP* (PDBID 7RHZ[14]), finalmente, dos de estos dímeros interactúan entre sí para ensamblar el 4) complejo Cre_4-*loxP* (intasoma, PDBID 1Q3U[13]).

Entre las interacciones Cre-DNA que estabilizan al complejo, se encuentran las hélices B y D que interactúan con el DNA en la región cercana al sitio de corte de *loxP* (al borde de la secuencia espaciadora); mientras que, la hélice J del dominio CAT, interactúa con el surco mayor a la mitad de uno de los RBE[2,3] (Figura 6A). Estudios previos han mostrado que residuos lejanos al sitio de corte también son importantes en el reconocimiento del sitio *loxP*. Tal es el caso del glutamato 262, (localizado en la hélice J) el cual ha mostrado ser importante para mantener la selectividad hacia *loxP* y discriminar otros sitios[16]. Esto se observó en mutantes hechas en el RBE, en las cuales, al mutar dicho residuo (E262) a cualquier otro residuo, la eficiencia de recombinación aumentó, permitiendo recombinar sitios *loxP* alterados y el sitio *loxP* silvestre. El uso de una amplia colección de mutantes tanto en Cre como en *loxP* han dado información sobre la selectividad de Cre. Estas mutaciones han sido descritas para residuos del sitio activo, regiones en el RBE o en el sitio de corte, así como en regiones alejadas de dichos sitios[17,18] (Figura 6B). Aunado a esto también las interacciones mediadas por agua son importantes en la lectura indirecta del ADN, lo que hace de la interfaz de Cre-*loxP* una red compleja de interacciones[19,16].

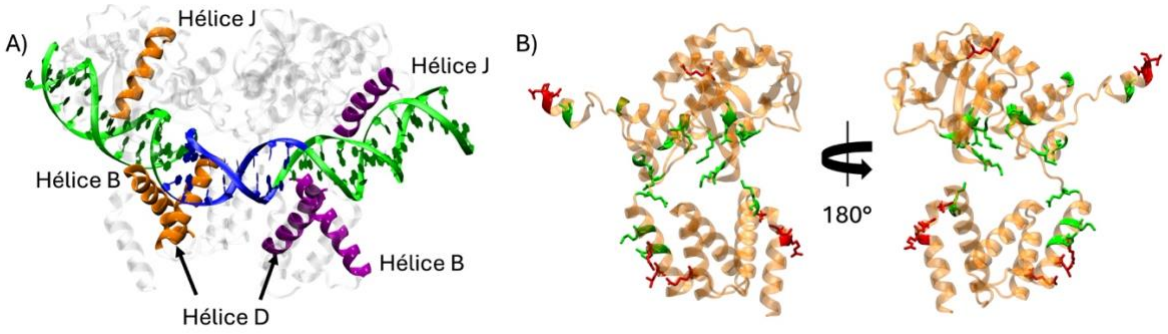


Figura 6. Interacciones Cre-*loxP* y mutantes de Cre. A) Interacción de las hélices B, D y J con *loxP* en el RBE izquierdo (hélices anaranjadas) y derecho (hélices moradas) marcadas sobre medio intasoma obtenido de la estructura 1Q3U. B) Mutantes descritas para la proteína Cre (PDBID 1Q3U cadena F). En rojo se muestran las mutantes que afectan la interfaz proteína-proteína. Las mutantes que afectan la eficiencia de recombinación se muestran en verde.

En el caso de las interacciones proteína-proteína, se identifican dos interfaces principales. La primera sucede entre dominios CBD, en la cual la hélice A interactúa con las hélices C y E de la subunidad vecina (Figura 7A). La segunda es entre dominios CAT[20], la cual consiste en una cavidad formada por las hélices F, I, K y L en la que se inserta la hélice N del monómero adyacente (Figura 7B). Ambas interfaces son importantes para la estabilidad del intasoma, ya que estudios previos han mostrado que la disrupción de éstas lleva a una pérdida de actividad en la reacción de recombinación[11,21,22]. El uso de mutantes que interfieren con la formación de interfaces proteína-proteína permitió la obtención de estructuras de un monómero unido a medio sitio *loxP* (PDB ID 7RHY) y un dímero unido a un sitio *loxP* invertido (*loxI*) completo (PDB ID 7RHZ)[14] (Figura 6B, residuos en rojo).

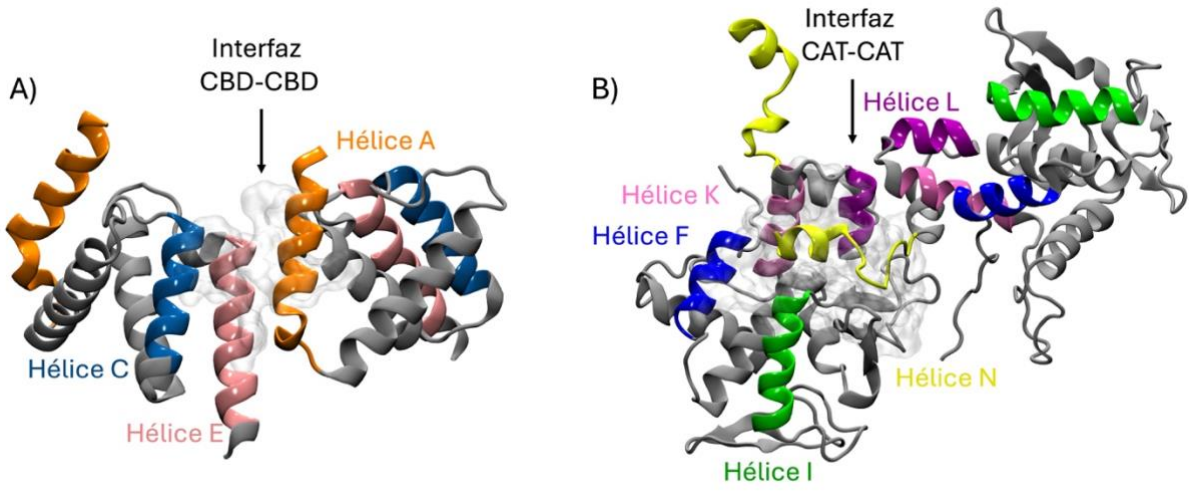


Figura 7. Interfaces proteína-proteína. A) Interfaz CBD-CBD, en anaranjado se muestra la hélice A, las hélices C y E que interactúan con esta hélice se muestran en azul y rosa, respectivamente. B) Interfaz CAT-CAT, se muestran las hélices F (azul), I (verde), K (rosa) y L (morado) que forman la cavidad en la que interactúa la hélice N (amarillo) del monómero adyacente. En ambos paneles la región de interacción se muestra como una superficie transparente.

1.4 Métodos computacionales

El estudio estructural de Cre ha sido desafiante debido a que, desde el armado del complejo, la reacción está sesgada hacia la formación del intasoma[14]. Lo mismo sucede con el estudio del sitio activo, en el cual se ha cambiado el fósforo del grupo fosfato por vanadio o azufre, esto con el fin de evitar que la reacción suceda[12,13]. En este escenario, los métodos computacionales han permitido el estudio de estos sistemas utilizando la versión silvestre tanto de Cre como del ADN.

Aunado a esto, los métodos computacionales dan un vistazo a nivel atómico a las interacciones que ocurren en el sistema; estas escalas son en su mayoría de difícil acceso mediante métodos experimentales[23]. Herramientas como acoplamiento molecular, dinámica esencial (análisis de componentes principales, análisis de modos normales, dinámica elástica (ENM), entre otros), y dinámica molecular son los métodos más populares debido a que nos brindan información estructural y

dinámica que puede ser comparable con datos experimentales. En este trabajo nos enfocaremos en las propiedades estructurales y dinámicas de Cre utilizando dinámica molecular.

1.4.1 Generalidades de la dinámica molecular

La dinámica molecular se ha posicionado como una de las herramientas más utilizadas para el estudio de proteínas y ácidos nucleicos que nos permite estudiar el comportamiento de sistemas biológicos en función del tiempo. La dinámica molecular s e basa en la resolución de las ecuaciones de movimiento de Newton (**F**=m**a**) para cada átomo del sistema[23].

Además, se cuentan con parámetros específicos que describen las interacciones entre átomos, los campos de fuerza. Estos campos de fuerza toman en cuenta longitud de enlace, ángulos, ángulos diedros, ángulos diedros impropios e interacciones no covalentes como las fuerzas de Van der Waals y electrostáticas. Entre los campos de fuerza más utilizados se encuentran CHARMM (CHemistry At Harvard Molecular Mechanics)[24], AMBER[25], GROMOS[26], y OPLS[27], entre otros. Todos estos campos de fuerza están parametrizados con datos experimentales que buscan describir de la mejor manera lo que sucede en el mundo real o en el tubo de ensayo.

Dado que las ecuaciones de movimiento se resuelven para cada átomo del sistema, una de las principales limitaciones de esta herramienta es el costo computacional. A mayor número de átomos, mayor el número de interacciones para calcular, lo cual resulta en una gran inversión de tiempo para obtener una simulación. Esto ha llevado al desarrollo de metodologías de muestreo acelerado, que permiten una exploración más rápida del espacio conformacional sin perder mucha información. Estrategias como intercambio de réplica, grano grueso y metadinámica son las más utilizadas[28,29].

Además de las estrategias de muestreo acelerado, se han desarrollado supercomputadoras que hacen uso de procesadores (CPUs) y tarjetas gráficas

(GPUs) que funcionan en sinergia para hacer dinámica molecular. El uso conjunto de ambas estrategias ha permitido la simulación de sistemas grandes tales como células completas, aerosoles, bicapas, y grandes complejos proteicos[30,31], entre otros.

2. Antecedentes

2.1.1 Estudios estructurales del complejo Cre-*loxP*

La información estructural disponible hasta el momento proviene principalmente de complejos tetraméricos, en su mayoría obtenidos por cristalografía de rayos X[2]. Además, también se cuenta con tres estructuras obtenidas mediante criomicroscopía electrónica. Estos complejos corresponden a un monómero de Cre unido a medio sitio *loxP* (PDB ID: 7RHY[14]), un dímero unido a un sitio *loxP* invertido (*loxI*) completo (PDB ID: 7RHZ[14]) y un intasoma con dos sitios *loxP* (PDB ID: 7RHX[14]).

Para la obtención de todas estas estructuras, se han empleado diversas estrategias que han permitido capturar complejos en distintas etapas de la reacción de recombinación. Además del uso de mutantes de azufre y vanadio para capturar al complejo covalente, también se utilizó un "sustrato suicida", el cual consistió en una secuencia de ADN a la que le hace falta un enlace fosfodiéster entre dos bases. Esto lleva a la formación de un *gap* después de formado el enlace covalente, ya que la base donadora del fosfato no está unida al resto de la secuencia de ADN, impidiendo el subsequente ataque al enlace fosfotirosina y, por tanto, la continuación de la reacción de recombinación[32] (Figura 8). Por otro lado, para capturar el intermediario de Holliday se han empleado mutaciones en los residuos del sitio activo que impiden la progresión de la reacción de recombinación, así como el uso de péptidos inhibitorios[33]. Todas las estructuras depositadas en la base de datos del PDB, así como su resolución y las mutaciones tanto en la proteína como en el ADN al que están unidas, se resumen en la Tabla 2.

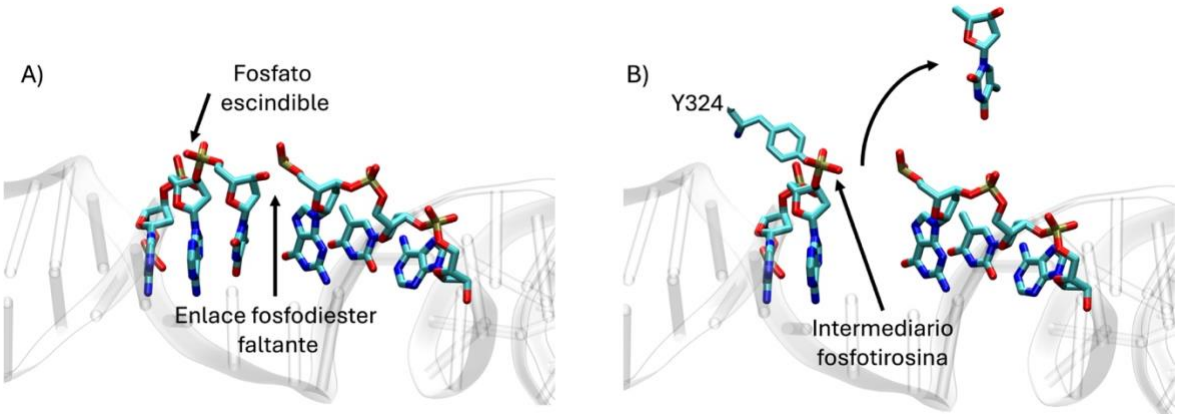


Figura 8. Sustrato suicida usado para captura el intermediario covalente de Cre. A) secuencia de ADN indicando la posición del fosfato escindible y el enlace fosfodiéster faltante (sustrato suicida). B) Formación del enlace covalente fosfotirosina entre la Y324 y el fosfato escindible. La base donadora del fosfato se libera llevando a la formación del *gap* y el intermediario covalente.

Tabla 2. Estructuras de Cre disponibles en el PDB. Para cada estructura depositada en el PDB se muestra la resolución de la estructura, las mutaciones o regiones faltantes en la proteína y las modificaciones en el ADN. La secuencia espaciadora se muestra subrayada en negro, las mutaciones y sustituciones en rojo y las regiones en las que se forma el intermediario covalente, en verde.

PDB ID	Resolución (Å)	Características	ADN
1CRX[32]	2.40	Intermediario covalente, *loxA*, residuos 20-341	TATAACTTCGTATAGCATATGCTATACGAAGTTAT
1DRG[34]	2.55	Trímero unido a DNA, residuos 180-188 faltantes, mutación Y324F, *loxA*	TATAACTTCGTATAGC

1F44[34]	2.05	Trímero unido a DNA, residuos 181-188 faltantes, mutación Y324F, *loxA*	TATAACTTCGTATAGC
1KBU[35]	2.20	Intermediario de Holliday, residuos 329-332 faltantes, *loxP*	ATAAGTTCGTATAATGTATGCTATACGAAGTTAT
1MA7[36]	2.30	Intasoma unido a mutante de *loxP*, residuos 328-332 faltantes	ATAACTTAGTATAATGTATGCTATACTAAGTTAT
1NZB[13]	3.10	Residuos 20-341, *loxP* con U= 5-Iodo-desoxiuridina	CGATAACUTCGTATAATGTATGCTATACGAAGTTATC
1OUQ[13]	3.20	Intermediario covalente, residuos 20-341; *loxP* con U= 5-Iodo-desoxiuridina y Fósforotioato	CGATAACUTCGTATAATGTATGCTATACGAAGTTATC
1PVP[19]	2.35	Intermediario de Holliday, residuos 20-341, mutante ALSHG en posiciones 174, 258, 259, 262, 266, residuos 327-331 faltantes, *loxM*	ATAACTCTATATAATGTATGCTATATAGAGTTAT
1PVQ[19]	2.75	Intermediario de Holliday, residuos 20-341, mutante LNSGG en posiciones 174, 258, 259, 262, 266, residuos 327-331 faltantes, *loxM*	ATAACTCTATATAATGTATGCTATATAGAGTTAT
1PVR[19]	2.65	Intermediario de Holliday, residuos 20-341, mutante LNSGG en posiciones 174,	ATAACTTCGTATAATGTATGCTATACGAAGTTAT

		258, 259, 262, 266, *loxP*	
1Q3U[13]	2.90	Residuos 20-341, *loxP* con U= 5-Iodo-desoxiuridina	CGATAACUTCGTATAATGTATGCTATACGAAGTTATC
1Q3V[13]	2.91	Intermediario covalente Residuos 20-341, *loxP* con U= 5-Iodo-desoxiuridina y Fósforotioato	CGATAACUTCGTATAATGTATGCTATACGAAGTTATC
1XNS[33]	2.80	Intermediario de Holliday, residuos 20-341, péptido inhibitorio, *loxP*	TATAACTTCGTATAATGTATGCTATACGAAGTTAT
1XO0[33]	2.00	Intermediario de Holliday, residuos 20-341, mutante R173K, *loxP*	TATAACTTCGTATAATGTATGCTATACGAAGTTAT
2CRX[37]	2.50	Residuos 199-209, 276-282 faltantes en cadena A. Residuos 199-206, 327-332 faltantes en cadena B. *loxA*	TATAACTTCGTATAGCATATGCTATACGAAGTTAT
2HOF[38]	2.40	Residuos 20-341, mutante K201A, *loxI* Residuos 198-210 faltantes en cadena A Residuos 328-333 faltantes en cadena B	TATAACTTCGTATAGCATACATTATACGAACTTAT
2HOI[38]	2.60	Residuos 20-341, mutante K201A, *loxI*	TATAACTTCGTATAGCATACATTATACGAACTTAT
3C28	2.60	Residuos 20-341, *loxP*	ATAACTTCGTATAATGTATGCTATACGAAGTTAT

		Residuos 186-192, 197-211, 279 faltantes en cadena A Residuos 327-332, 334 faltantes en cadena B	
3C29	2.20	Residuos 20-341, K201 incompleta, *loxP* con diaminopurina	TATAACTTCGTATAATGTATGCTATACGAAGTTAT
3CRX[37]	2.50	Intermediario de Holliday, mutante R173K, *loxA*	TATAACTTCGTATAGCATATGCTATACGAAGTTAT
3MGV[12]	2.29	Residuos 20-341, *loxA*, Vanadato-Tyr	TATAACTTCGTATAGCATATGCTATACGAAGTTAT
4CRX[39]	2.20	Mutante R173K, *loxA*	TATAACTTCGTATAGCATATGCTATACGAAGTTAT
5CRX[39]	2.70	Residuos 194-209, 317-325 faltantes en cadena A Residuos 315-343 faltantes cadena B, mutante Y324F, *loxA*	TATAACTTCGTATAGCATATGCTATACGAAGTTAT
7RHX*[14]	3.23	Residuos 20-341, mutante K201A, *loxI*	GCATAACTTCGTATAGCATACATTATACGAAGTTATCG
7RHY*[14]	3.91	Residuos 20-329; horquilla *loxA*. Mutantes D33A, A36V, R192A	GCATAACTTCGTATAGCATATGC -GAA-
7RHZ*[14]	4.48	Residuos 20-341, *loxI* Residuos 199-207 faltantes en cadena A y mutantes D33A, A36V, R192A. Residuos 200-207, 328-341 faltantes en cadena B y mutantes	GCATAACTTCGTATAGCATACATTATACGAAGTTATCG

		R72E, K115D, R119D.	

Estas estructuras han dado información detallada sobre las interacciones presentes en el sitio activo, las interacciones encargadas de estabilizar al intasoma, así como el efecto de mutaciones sobre el ADN o la proteína en la eficiencia de recombinación y selección de la secuencia blanco.

Estudios biofísicos de Cre han demostrado que esta proteína es monomérica en ausencia de ADN[5.] Sin embargo, en presencia de este, el equilibrio está sesgado hacia la formación del intasoma[38]. Para los dominios por separado hay menos información estructural. Experimentos de RMN mostraron que ambos dominios son independientes y pueden ser estudiados por separado[11,40]; esto se determinó comparando los espectros de HSQC para la proteína completa y los dominios aislados, mostrando que los espectros sobrelapan entre sí, lo cual llevó a la conclusión que se pliegan igual en ambos contextos[11]. Este mismo trabajo realizó experimentos de PRE (Paramagnetic Relaxation Enhancement) para el dominio CAT aislado, llevando a la propuesta de un modelo en el que la hélice N puede interactuar con los residuos del sitio activo llevando a la formación de un complejo auto-inhibitorio. Sin embargo, a pesar de esta información estructural disponible, aún no existe una estructura experimental para este complejo. En el caso del dominio CBD, no existe información experimental disponible, tanto para el dominio CBD aislado, ni para este dominio unido a ADN. Sin embargo, estructuras cristalográficas de los complejos tetraméricos muestran que los primeros 19 residuos del extremo N-terminal son intrínsecamente desordenados ya que no se han visto en ninguna de las estructuras disponibles, pero se ha propuesto que estos residuos pueden interactuar en una cavidad hidrofóbica presente entre las hélices A y B[41]. Curiosamente, la reacción de recombinación no se ve afectada en ausencia de hasta 20 residuos de este extremo[42]. Finalmente, las estructuras obtenidas por criomicroscopía electrónica para un dímero y un monómero unido a ADN, muestran

una posición diferente de la hélice A comparada a su posición vista en un intasoma[14] (Figura 9); esta posición tipo monómero/dímero también es vista en las estructuras de un trímero[34]. Además, en el contexto de un intasoma, el ADN tiene un doblez de aproximadamente 90° comparado con el visto en un dímero, el cual es de aproximadamente 50°[14]. Esto sugiere que se necesitan interacciones cuaternarias para doblar al ADN y para colocar a dicha la hélice A en la posición adecuada para llevar a cabo la reacción de recombinación.

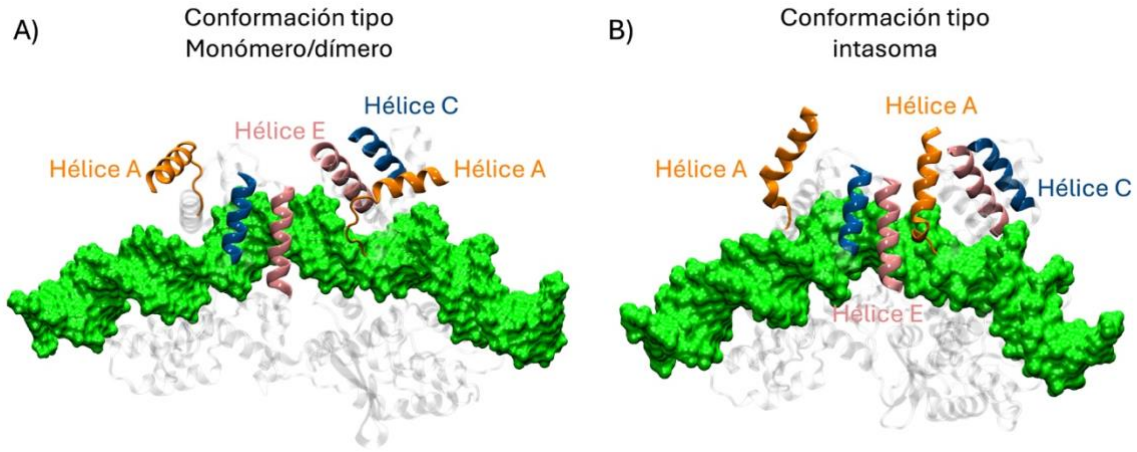


Figura 9. Diferentes posiciones de la hélice A. A) Posición de la hélice A en un dímero y B) en un intasoma. En ambos casos las hélices A, C y E se muestran en anaranjado, azul y rosa, respectivamente, mientras que *loxP* está en color verde.

2.1.2 Dinámica molecular del sistema Cre-*loxP*

Debido a su relevancia en el ámbito biotecnológico, se han realizado simulaciones de dinámica molecular para estudiar el extremo N-terminal de Cre en el contexto del intasoma y su papel en la reacción de recombinación[41]. Por otra parte, se han hecho estudios con mutantes de *loxP* para evaluar la afinidad de Cre hacia diferentes secuencias[18], y también se han evaluado mutantes de Cre que incrementan el número de interacciones hidrofóbicas y su efecto sobre el reconocimiento de *loxP*[16,43].

El estudio de Cre-*loxP* mediante el uso de herramientas computacionales, es un campo abierto que permite la exploración de este sistema en múltiples contextos. Además, como se mencionó anteriormente, la flexibilidad intrínseca de Cre en ausencia de interacciones proteína-proteína y proteína-ADN dificulta su caracterización experimental. Esto vuelve interesante el uso de dinámica molecular para contestar y entender mejor los mecanismos que siguen sin ser resueltos para este sistema, tales como la dinámica de Cre libre, la selección de la secuencia blanco y los cambios conformacionales asociados a la unión y reconocimiento, entre otros.

3. Justificación

A pesar de ser un sistema ampliamente utilizado en ingeniería genética, y de conocerse los cambios estructurales que ocurren durante la reacción de recombinación, aún no existe información estructural sobre Cre en ausencia de ADN. Conocer esta estructura, así como los cambios conformacionales asociados a los primeros pasos de formación del intasoma, nos permitirá entender mejor el mecanismo de ensamblaje del complejo y a proponer nuevas estrategias que aumenten la especificidad y eficiencia de este sistema. Además, se desconoce el papel de la hélice N en la determinación de un monómero activo y en el doblez de *loxP.*

4. Hipótesis

La posición de la hélice N tiene un efecto sobre la dirección del doblez de *loxP,* este doblez determinar al monómero activo.

La estructura de Cre unida a *loxP,* se encuentra en una conformación “restringida”, esta estructura está estabilizada por múltiples interacciones proteína-proteína y proteína-ADN. En ausencia de estas interacciones Cre visita múltiples

conformaciones promoviendo un mecanismo de selección de conformación para el armado del intasoma. Los dominios por separado son independientes y muestran un comportamiento similar al observado en la proteína completa.

5. Objetivo

Caracterizar el espacio conformacional de un dímero de Cre unidos a *loxI*, asi como de Cre en ausencia de interacciones cuaternarias y de sus dominios individuales en ausencia de ADN.

5.1 Objetivos particulares

- Caracterizar el espacio conformacional de un dímero de Cre unido a un sitio *loxI*.
- Caracterizar el espacio conformacional de Cre monomérica en ausencia de interacciones proteína-proteína y proteína-ADN.
- Caracterizar el espacio conformacional del Dominio CBD en ausencia de ADN.
- Caracterizar el espacio conformacional del dominio CBD unido a la horquilla *loxA* (CBD-*loxA*).
- Caracterizar el espacio conformacional del dominio CAT en ausencia de ADN y de la hélice N (CAT-delN).

6. Metodología

6.1 Obtención de estructuras iniciales

Todas las estructuras experimentales utilizadas en este estudio fueron obtenidas de la base de datos del Protein Data Bank[44] (PDB: https://www.rcsb.org/).

En la mayoría de las estructuras experimentales disponibles los residuos 1 a 19 no están presentes, esto debido a que dichos residuos no han sido resueltos o tienen baja densidad electrónica, y como se mencionó anteriormente, la eliminación de

estos residuos no tiene efecto sobre la reacción de recombinación y su unión al ADN[42]. Por otro lado, los residuos 342 - 343 tampoco están presentes en dichas estructuras.

La mayoría de las estructuras experimentales tienen residuos faltantes que no pudieron ser resueltos (Tabla 2), o algún tipo de mutación tanto en la proteína como en el ADN al que están unidas. La estructura con código de PDB 1Q3U es la única estructura que no tiene mutaciones en la proteína y solo tiene una sustitución por 5'-desoxiuridina en la posición 11 (letra U marcada en rojo en la secuencia). A pesar de ser una estructura con una resolución de 2.9 Å, todos los residuos pudieron ser modelados y no tiene regiones no resueltas. Por esta razón se tomó esta estructura para obtener las coordenadas del dominio CBD (residuos 20 al 128), CAT-delN (residuos 138 a 326), CAT (residuos 128 al 341), y los sistemas activo e inactivo de Cre completa (residuos 20 a 341). Por otro lado, sólo existe una estructura para el monómero y dímero unidos al ADN; estas estructuras fueron seleccionadas y se modelaron las regiones faltantes, así como la reversión a silvestre de las mutaciones. Al igual que en los sistemas obtenidos de 1Q3U, se conservaron los mismos grupos de residuos para CBD, CAT y Cre libre. Finalmente, para los sistemas dímero 7RHZ se tomó la estructura proveniente del PDB y se modelaron las regiones faltantes; para el sistema CBD-*loxA* se conservaron las mismas coordenadas que el CBD aislado, así como las de la horquilla *loxA* tomada de 7RHY. La Tabla 3 muestra el código PDB de cada una de las estructuras utilizadas y los sistemas obtenidos a partir de ellas.

Tabla 3. Colección de estructuras iniciales. Para cada código de PDB se muestran los sistemas obtenidos que fueron simulados en este proyecto.

Sistema **PDBID**	**CBD, CAT**	**CBD *loxA***	**CAT delN**	**Cre libre**	**Dímero 7RHZ**

1Q3U:A (monómero inactivo de un intasoma)	X		X	X	
1Q3U:F (monómero activo de un intasoma)	X		X	X	
7RHY (monómero unido a medio sitio *lox*)	X	X		X	
7RHZ:Yin (cadena A del dímero unido a *loxI*)	X			X	X
7RHZ:Yang (cadena B del dímero unido a *loxI*)	X			X	

6.2 Optimización de estructuras

La optimización de las estructuras se realizó en múltiples etapas. Para 7RHY y 7RHZ, provenientes de criomicroscopía electrónica (CryoEM), las cuales no presentan un enlace cis-peptídico entre los residuos 64 y 65, se corrigió dicho enlace ajustando el ángulo diedro formado entre ambos residuos utilizando VMD[45]. Los residuos 198 a 208 de ambos monómeros de 7RHZ, así como del monómero 7RHY no fueron resueltos en la estructura experimental. Para su reconstrucción se usó el servidor Swiss-Model[46], tomando como molde la estructura 1Q3U:A,

correspondiente al monómero inactivo proveniente de un intasoma. Por otro lado, los residuos 330 a 341 de la estructura 7RHY y los residuos 328 a 341 de la cadena Yang de 7RHZ tampoco fueron resueltos. Estos residuos fueron insertados en las respectivas estructuras copiando las coordenadas de la estructura 7RHZ:Yin mediante la técnica de grafting, la cual consiste en copiar las coordenadas de una estructura y pegarlas en la estructura de interés.

Una vez obtenidas las estructuras completas, se añadieron los átomos de hidrógeno con el servidor CHARMM-GUI[47]. Debido que no se sabe la estructura de los primero 19 residuos del extremo N-terminal y los últimos dos del C-terminal, se usaron los parches ACE y CT2 para acetilar y amidar los extremos N- y C-terminales[48], respectivamente, esto con el fin de evitar el efecto de cargas extra en los extremos.

Como se mencionó anteriormente, la estrategia empleada para obtener la estructura del monómero y el dímero unidos a ADN consistió en utilizar mutantes que interfirieran con las interfaces proteína-proteína, evitando así la formación del intasoma. Estas mutaciones, presentes en las estructuras 7RHY (D33A/A36V/R192A) y 7RHZ (Yin: D33A/A36V/R192A; Yang: R72E/L115D/R119D), fueron revertidas a su versión silvestre. Además, para los sistemas unidos a ADN (CBD-*loxA* y dímero 7RHZ) se utilizaron los parches 5TER, 3TER en los extremos 5’ y 3’ del ADN, respectivamente[48]. Todas estas modificaciones se realizaron utilizando el mismo servidor.

Previo a la construcción de las celdas de simulación, todas las estructuras obtenidas en el paso anterior fueron minimizadas con el programa CHARMM48[24] siguiendo la serie de pasos que se muestra a continuación:

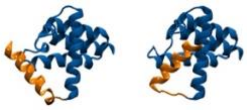
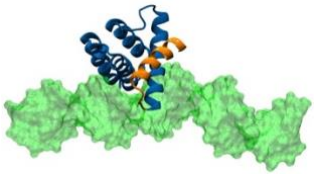
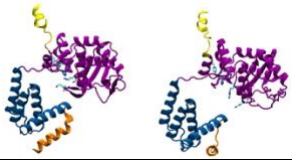
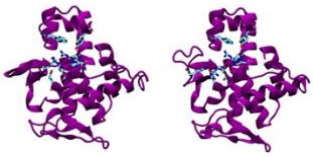
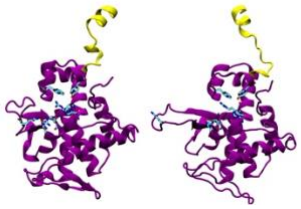
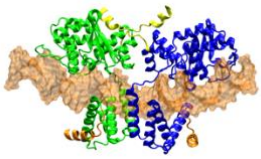
1. 100 pasos de minimización de energía usando el algoritmo steepest descent para el enlace peptídico en *cis* entre los residuos 64 y 65 para los sistemas provenientes de 7RHY y 7RHZ.

2. 100 pasos de minimización de energía para el enlace entre los residuos 330-331 y 328-329 usando el algoritmo steepest descent para los sistemas provenientes de las estructuras 7RHY cadena Yang de 7RHZ, respectivamente.
3. 100 pasos de minimización de energía para todos los átomos de hidrógeno usando el algoritmo steepest descent para todos los sistemas.
4. 100 pasos de minimización de energía para todo el sistema usando el algoritmo conjugate gradient para todos los sistemas.

6.3 Construcción de celdas de simulación

Todas las celdas fueron construidas en el servidor CHARMM-GUI[49], usando las estructuras optimizadas obtenidas en el paso previo. Cada estructura fue sumergida en una celda cúbica con el modelo de agua TIP3P. Para los sistemas CBD aislado y CAT-delN se utilizó una capa extra de agua de 15 Å entre la proteína y el borde de la celda para evitar que haya interacciones de la proteína con su imagen. Para los sistemas CBD-*loxA*, Cre completa y dímero unido a *loxI*, el tamaño de la caja se determinó midiendo previamente la estructura seca en la terminal TK de VMD. Para todos los sistemas se utilizó una concentración de sal de 0.15 M de NaCl; los iones fueron agregados con el método Monte Carlo. Se ajustó una temperatura constante de 303.15 K y se utilizó el algoritmo de repartición de masas (HMR)[50] así como la corrección para las interacciones entre residuos aromáticos (WYF)[51]. Posteriormente, se generaron los archivos necesarios para dinámica molecular con el campo de fuerza de CHARMM36m[52] y el software GROMACS[53]. La tabla 4 muestra el tamaño de cada uno de los sistemas.

Tabla 4. Tamaño de los sistemas obtenidos en el servidor CHARMM-GUI. Para cada sistema se muestra el tamaño de la caja, el número total de átomos, así como el número total de iones y de moléculas de agua. La última columna muestra la estructura de cada sistema: CBD (estructura tipo intasoma y tipo dímero), CBD-*loxA*, Cre libre (estructura tipo intasoma y tipo dímero), CAT-delN (monómero activo e inactivo), CAT delN (monómero activo e inactivo) y dímero 7RHZ.

	Tamaño de la caja (Å)	#Átomos	#Aguas	#Iones	Estructuras
CBD	82X82X82	51,871	16,652	Na: 47 Cl: 54	
CBD-*loxA*	118X83X75	68,492	21,652	Na: 102 Cl: 61	
Cre libre	115X115X115	143,358	45,985	Na: 130 Cl: 142	
CAT-delN	71X71X71	33,213	10,065	Na: 28 Cl: 35	
CAT	74X74X74	38,107	11,617	Na: 32 Cl: 36	
Dímero 7RHZ	155X105X87	131,808	39,619	Na: 161 Cl: 111	

Finalmente, para los sistemas unidos al ADN (CBD-*loxA* y el dímero 7RHZ), se asignaron restricciones armónicas al principio y al final del dúplex *loxI* en el dímero

7RHZ, y para la horquilla del sistema CBD-*loxA*, se aplicaron restricciones al principio de la horquilla y en el par previo a la vuelta GAA en la punta de la horquilla (Figura 10). En ambos sistemas todas las restricciones se aplicaron sobre los átomos de pares de base GC que forman puentes de hidrógeno (Figura 10B). Para que las "tapas" fueran consistentes en ambos sistemas, sólo se conservaron dos pares de bases GC en la punta de ambas secuencias.

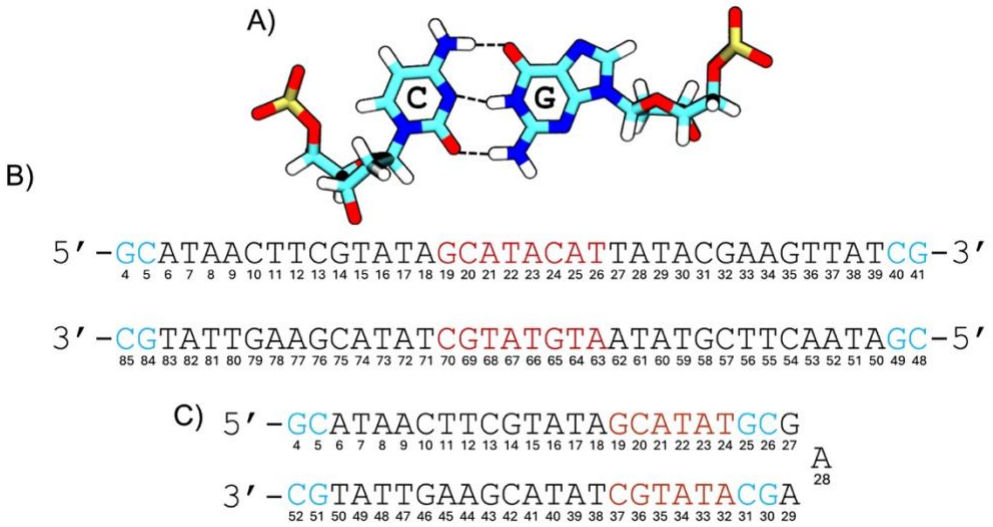


Figura 10. Secuencias *lox* y restricciones. A) Átomos entre los cuales se impusieron las restricciones. Todas tienen la misma fuerza de 800 kJ mol^{-1} nm^{-2}. B) Secuencia de *loxI* y C) horquilla *loxA*, debajo se muestra la numeración utilizada en cada caso. La secuencia espaciadora se muestra en rojo; en azul las "tapas" en los extremos 5' y 3' en las cuales se colocaron las restricciones armónicas.

6.4 Dinámica molecular

Todas las simulaciones se llevaron a cabo con el campo de fuerza de CHARMM36m y el ensamble NPT a 1 atmósfera de presión y 303.15 K de temperatura. Se realizó una minimización sin restricciones de 5,000 pasos con el algoritmo steepest descent para todo el sistema, seguida por dos equilibraciones. Las dos equilibraciones se realizaron con el barostato de Berendsen, debido a que

permite una convergencia más rápida de la energía del sistema, usando un paso de 2 fs de tiempo de integración. La primera equilibración fue de un nanosegundo para el solvente (agua y iones), aplicando restricciones de posición para la proteína y el ADN. En la segunda equilibración se aplicaron restricciones para la cadena principal de la proteína y del ADN (en el caso de los sistemas CBD-*loxA*, dímero 7RHZ), dejando así que se equilibraran el solvente y las cadenas laterales. Finalmente, se eliminaron todas las restricciones y se corrió un nanosegundo de dinámica molecular para todo el sistema. La producción se llevó a cabo con un paso de integración de 4 fs, y el integrador leap-frog verlet para resolver las ecuaciones de movimiento, ya que es un algoritmo preciso y estable que permite la simulación de sistemas complejos; la presión se controló isotrópicamente usando el baróstato de Parrinello-Rahman y la temperatura con el termostato de Nose-Hoover. La lista de vecinos se actualizó con una frecuencia de 20 pasos usando un corte de 12 Å. Las interacciones electrostáticas fueron tratadas usando el algoritmo Particle Mesh Ewald (PME) con un corte de 12 Å, se utilizó este algoritmo ya que permite un cálculo más preciso de las interacciones de corto y largo alcance. Las interacciones de Van de Waals fueron calculadas usando un corte de 12 Å y la función Force-switch empezando en 10 Å. Se usó el algoritmo LINCS[54] para todos los átomos de hidrógeno unidos a átomos pesados; las coordenadas fueron guardadas cada picosegundo. Todas las simulaciones fueron corridas en bloques de 1 ns en el programa GROMACS. La tabla 5 muestra el número de réplicas, tiempo acumulado de simulación y versión de GROMACS para cada sistema. Debido a que a lo largo del proyecto se utilizaron diferentes versiones de GROMACS, todas las simulaciones se llevaron a cabo siguiendo exactamente los mismos parámetros para todos los sistemas, esto con el fin de poder comparar propiedades y comportamientos. Además de que se respetó la versión del software utilizada desde el inicio para cada réplica.

Tabla 5. Tiempo acumulado de simulación para cada sistema.

Sistema	Nombre	Réplicas por sistema	Tiempo por réplica (μs)	Tiempo acumulado (μs)	Versión de GROMACS
Cre completa	1Q3U:F	4	5	100	2019.2
	1Q3U:A				2021.3
	7RHY				
	Yin				2023
	Yang				
CBD asilado y CAT	1Q3U:F	4	5	100	2021.3, 2023
	1Q3U:A				
	7RHY				
	Yin				
	Yang				

CBD-*loxA*	7RHY	4	5	20	2021.3
Dímero 7RHZ	7RHZ	4	5	20	2021.3
CAT-delN	1Q3U:F	3	5	30	2019.2
	1Q3U:A				

Debido a que la calidad de la estructura de 1Q3U no es óptima, obtuvimos una versión refinada de este cristal con el servidor PDB-REDO[55]. Este servidor permite generar estructuras optimizadas utilizando los datos experimentales de las entradas depositadas en el PDB, o la bien refinar estructuras propias. La estructura obtenida con PDB-REDO presenta mejores valores de R (0.2079 vs 0.2310) y R_{free} (0.2367 vs 0.2546) en comparación con la estructura original del PDB. A esta estructura se le añadieron los átomos de hidrógeno, y se realizó la minimización de energía como se indicó para las estructuras iniciales. Posteriormente, se realizó un cálculo de RMSD tomando todos los átomos excepto los átomos de hidrógeno de esta estructura contra la estructura optimizada. Los valores de RMSD obtenidos son de 0.44 Å para el monómero activo y de 0.43 Å para el monómero inactivo (Anexo II). Con esta comparación se confirma que la estructura de la que se partió para la simulación es de buena calidad. Además, a lo largo de la dinámica molecular, unos de los objetivos es muestrear todas las posibles posiciones de las cadenas laterales, así como cambios conformacionales en la proteína.

6.5 Cálculo de propiedades estructurales

Para cada sistema se calcularon diferentes propiedades estructurales para obtener información sobre su dinámica individual. En la sección de resultados se ahondará más en los detalles de cada uno de ellos. A continuación, se describe la

metodología para los cálculos que se comparten entre la mayoría de los sistemas. Además, todas las gráficas se realizaron con la librería Matplotlib[56] de Python

6.5.1 Radio de giro y RMSD

Se utilizaron las herramientas *gyrate* y *rms* en GROMACS para el cálculo de radio de giro y RMSD contra la estructura inicial. Ambos cálculos fueron hechos tomando en cuenta solo los Cα. En el caso del radio de giro para los residuos del sitio activo, se tomaron en cuenta todos los átomos de los residuos que conforman a la constelación catalítica (R173, E176, K201, H289, R292, W315, Y324).

6.5.2 Ángulos entre hélices

Se identificaron varios pares de hélices alfa en ambos dominios que pueden dar información sobre el comportamiento de Cre en diferentes contextos. Para los sistemas que cuentan con el dominio CBD (CBD aislado, CBD-*loxA*, Cre completa, dímero 7RHZ), se tomaron las hélices A - B y se calculó el ángulo y la distancia entre ellas usando el programa CHARMM48 con el algoritmo de Chotia et. Al[57]. Esto debido a que este algoritmo es más eficiente para hélices que están en contacto en comparación con hélices que están alejadas entre sí. Para el resto de los ángulos y distancias entre las hélices seleccionadas se usó la herramienta *gangle* en GROMACS, tomando el primer y último Cα, así como el Cα central de cada hélice para el cálculo del ángulo diedro y la distancia, respectivamente. Para los sistemas con el dominio CAT (Cre completa, dímero 7RHZ y CAT-delN) se tomaron las hélices I – K y para los sistemas con ambos dominios (Cre completa y dímero 7RHZ) se tomaron las hélices D – G, además de las hélices de los dominios individuales.

6.5.3 Contactos entre átomos pesados

Todas las interacciones proteína-proteína y proteína-ADN fueron calculadas usando el programa CHARMM48, tomando en cuenta solo las interacciones entre átomos pesados (sin hidrógenos), y un corte de 6 Å. Interacciones específicas entre diferentes regiones de la proteína y/o ADN fueron filtradas de esta lista general. El

número total de interacciones fue calculado como el promedio de las cuatro réplicas para cada sistema.

6.5.4 Estructura secundaria

La estructura secundaría se calculó con DSSP[58] dentro de GROMACS2023 para todos los residuos de cada sistema. Se sumaron todos los valores de hélices α, 3_{10} y π para obtener el total de estructura tipo hélice; del mismo modo, se agruparon los valores hoja β y puentes β para obtener el total de estructura tipo β, esto de acuerdo con la clasificación hecha en la base de datos UNIPROT[59].

7. Resultados

7.1 Dímero 7RHZ

7.1.1 Armado del complejo y validación experimental

Como se mencionó anteriormente, las partes faltantes del complejo dimérico unido a la secuencia *loxI* se modelaron con el servidor Swiss Model tomando como molde el monómero inactivo del complejo 1Q3U:A, seguido por la reversión de las mutaciones para obtener la versión silvestre de la proteína; además, las coordenadas de la hélice N fueron pegadas en esta estructura copiando las coordenadas de la estructura 7RHZ:Yin. La Figura 11 muestra la estructura inicial obtenida después de estas modificaciones.

En dicha estructura la hélice N del monómero Cre^{Yin} (azul) está invadiendo la cavidad del monómero Cre^{Yang} (rojo), y la horquilla β2/3 en la que se encuentra la K201 está lejos del ADN. Esto último puede estar dado por la utilización de la cadena inactiva de 1Q3U como molde. Ya que esta es una región definida como intrínsecamente desordenada, se espera que esta región muestree múltiples conformaciones. Además, la carga negativa del ADN también podría tener un efecto sobre la posición de este residuo acercándolo hacia el sitio activo.

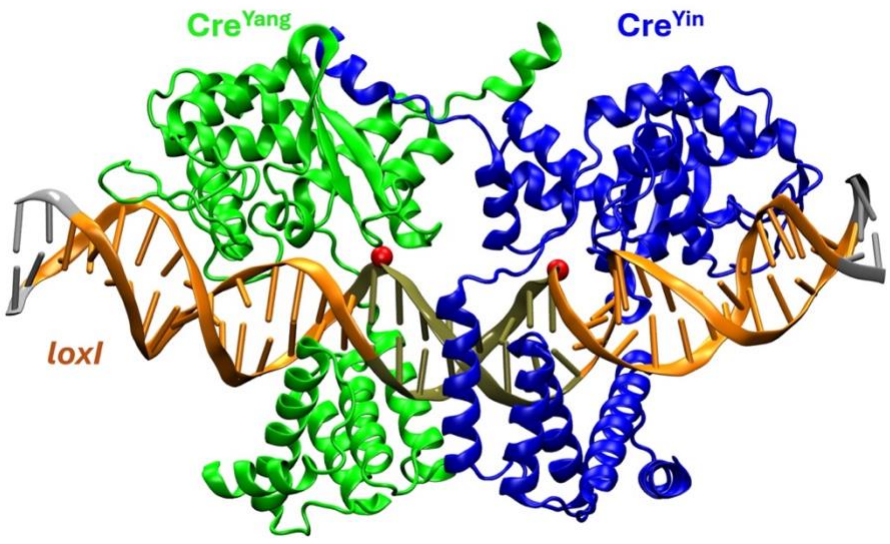


Figura 11. Estructura del dímero 7RHZ completo. Cada monómero se muestra en diferente color (Cre^{Yin} en azul, Cre^{Yang} en verde), la secuencia espaciadora se muestra en café; mientras que, los RBE están en anaranjado y las regiones con restricciones, en gris. La posición del fosfato escindible se marca en rojo sobre la secuencia espaciadora.

Para validar esta construcción se calcularon los contactos entre los átomos pesados de cada monómero y los átomos del azúcar del ADN. Estos contactos se compararon con el patrón de protección contra el radical OH^- obtenido mediante footprinting[60]. El objetivo de este experimento, en el contexto de un complejo proteína-ADN, es identificar las regiones del ADN donde se une la proteína, ya que dicha unión impide el corte del ADN promovido por el radical OH^-. Los resultados muestran que, en general, las regiones del ADN que están en contacto con cada monómero de Cre a lo largo de la simulación, correlacionan bien con las regiones obtenidas experimentalmente (bases marcadas en rosa en la Figura 12). Por otro lado, las regiones que corresponden a un bajo número de contactos también correlacionan con las regiones en las que hubo mayor frecuencia de corte por el radical OH- (regiones no protegidas). Las diferencias observadas entre los valores experimentales y los obtenidos en la simulación podrían explicarse por una diferencia en la curvatura entre un sitio *loxP* en el contexto de un tetrámero (mayor curvatura) y la de un dímero (menor curvatura). Este parámetro podría tener un efecto sobre las regiones que cubre cada monómero según la curvatura, ya sea cubriendo o descubriendo diferentes regiones. Con esta validación se confirma que el complejo es estable y se mantiene unido a las regiones determinadas experimentalmente, lo que nos permite buscar interacciones específicas entre residuos de Cre y las bases de *loxI*.

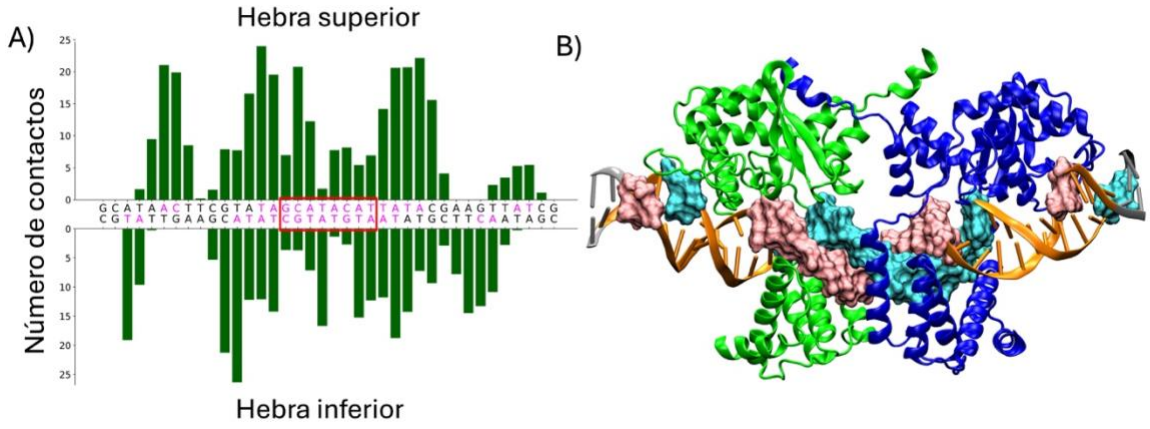


Figura 12. Comparación del patrón de protección experimental con los contactos proteína-ADN. A) Gráfica de barras que muestra el promedio de contactos entre la proteína y el azúcar de cada base (arriba: hebra superior; abajo hebra inferior). La secuencia de *loxI* se muestra en medio de ambas gráficas con la secuencia espaciadora encerrada en rojo, en rosa se marcan las regiones en las que hay protección contra el radical OH^- experimentalmente. B) Estructura del dímero 7RHZ mostrando en superficie las regiones protegidas por cada monómero (rosa: hebra superior; cian: hebra inferior).

7.1.2 Formación y ruptura de interfaces

En el complejo del intasoma, Cre realiza múltiples interacciones proteína-proteína, y todas estas interacciones tienen un papel clave en la estabilidad del complejo. Estas interacciones suceden principalmente entre CBD-CBD y CAT-CAT[3]. En particular, en el contexto del dímero, la hélice N de Cre^{Yin} interactúa en la cavidad de Cre^{Yang} dejando a la hélice N de Cre^{Yang} sin una cavidad vecina con la cual interactuar. Para evaluar la estabilidad del complejo y la formación o ruptura de estas interfaces, se calculó el número total de contactos entre átomos pesados realizados entre ambos monómeros (tomando en cuenta ambos dominios) y de los residuos del extremo C-terminal (residuos 326-341) con la cavidad del monómero adyacente. En ambos casos, los resultados se compararon con el número total de interacciones de una interfaz tipo dímero e intasoma. La Figura 13 muestra el número de contactos entre monómeros en cada réplica, en todos los casos, estas

interacciones aumentan en comparación con las observadas en la estructura inicial (línea continua en Figura 13A). Sin embargo, en ninguno de los casos el número de contactos alcanza al que hay en un complejo tipo intasoma (línea punteada en Figura 13A). Esto se observa con mayor claridad en la Figura 13D, que representa el número máximo de contactos entre subunidades. A pesar del incremento en los contactos entre dominios CBD con respecto a la estructura inicial (Figura 13B), estos siguen siendo pocos comparados con los del dímero proveniente de un intasoma (Figura 13C).

Además, la Figura 13D también muestra que la mayoría de los contactos se establecen entre la hélice N y la cavidad de la subunidad adyacente, lo que favorece la formación de un dímero en el que cada hélice N invade a la cavidad vecina. Estas interacciones recíprocas solo se observan en el contexto del dímero, ya que en el intasoma cada hélice N interactúa con un monómero adyacente, conectando así a los cuatro monómeros que conforman el complejo tetramérico.

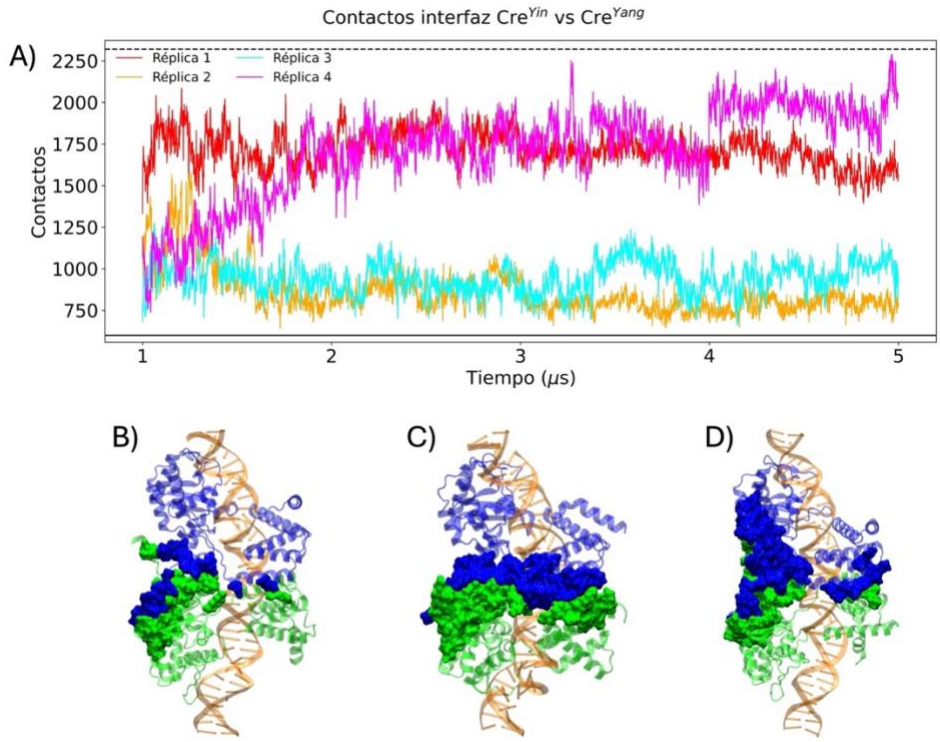


Figura 13. Formación y rompimiento de interfaces. A) Número total de contactos entre monómeros a lo largo del tiempo para cada réplica. Las líneas punteadas

representan el número de contactos iniciales para el intasoma; mientras que, la línea continua marca los iniciales del dímero. Los residuos de Cre^{Yin} (azul) y Cre^{Yang} (verde) de la interfaz están marcados en superficie para B) dímero (7RHZ), C) intasoma (1Q3U) y D) máximo número de contactos entre dominios encontrado durante la simulación.

La comparación de la Figura 13C con las Figuras 13B y 13D, muestra que la interfaz del intasoma tiene asociada una mayor curvatura en el ADN, la cual, parece promover interacciones CBD-CBD y CAT-CAT (Figura 13). Sin embargo, esta conformación de *loxI* no se ha logrado alcanzar en el tiempo de simulación. Esto podría sugerir que la conformación funcional del ADN en el intasoma es alcanzable sólo en el contexto del complejo tetramérico.

Por otro lado, los contactos entre la región C-terminal de Cre y la cavidad adyacente muestra un comportamiento dinámico, en el cual este segmento puede tener tres posibles conformaciones, las cuales se describen a continuación y se muestran en la Figura 14:

1.- C-terminal de Cre^{Yin} se queda en la cavidad de Cre^{Yang} mientras el C-terminal de Cre^{Yang} adquiere múltiples conformaciones en múltiples posiciones (Figura 14B, panel 1).

2.- C-terminal de Cre^{Yin} se queda en la cavidad de Cre^{Yang} mientras que el C-terminal de Cre^{Yang} entra a la cavidad de Cre^{Yin} formando un complejo en el cual ambas cavidades están ocupadas (Figura 14B, panel 2).

3.- Intercambio de segmentos, en el que el C-terminal de Cre^{Yin} sale de la cavidad de Cre^{Yang} mientras que el C-terminal de Cre^{Yang} entra a la cavidad de Cre^{Yin} (Figura 14B, panel 3).

La posición de este segmento podría tener un efecto sobre el doblez del ADN, ya que estudios previos[6] sugieren que la posición de la hélice N determina qué monómero de Cre estará en conformación activa y, por ende, iniciará la reacción de recombinación.

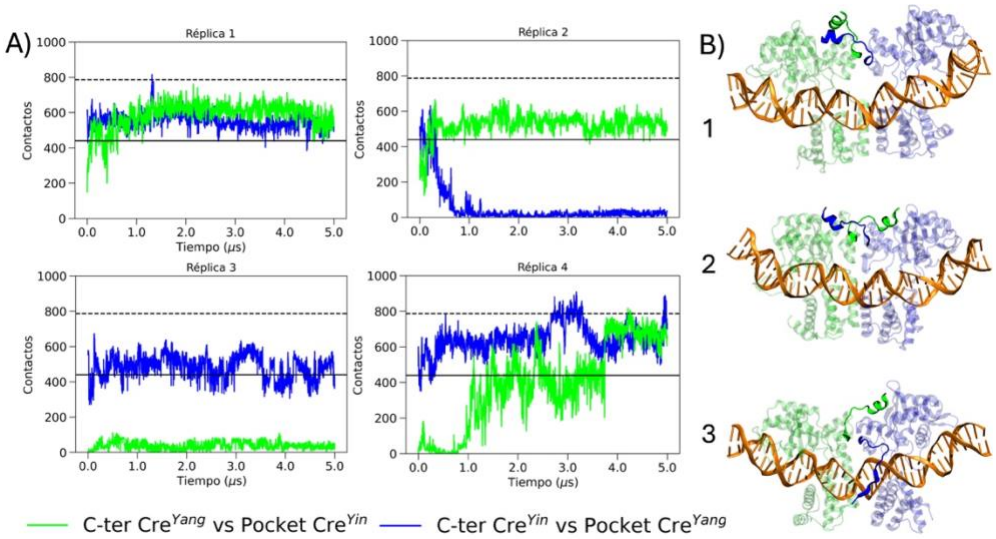


Figura 14. Invasión de la hélice N hacia el monómero adyacente. A) Número total de contactos entre el extremo C-terminal de cada monómero y la cavidad del monómero adyacente (azul, yin; verde, yang). La línea punteada representa el número de interacciones en un intasoma; mientras que la línea continua representa aquellas en un dímero. B) Estructura de cada caso descrito; 1) las hélices terminan como iniciaron, 2) ambas hélices ocupan la cavidad vecina, 3) hay intercambio de hélices invasoras.

7.1.3 Dinámica de *loxI* y la posición de la hélice N

Un proceso clave para que la reacción de recombinación ocurra es la determinación de un monómero activo. Durante la unión inicial de Cre al sitio *loxP,* se ha propuesto que la primera hélice N que invade al monómero adyacente define cuál será este monómero[6], ya que un monómero activo tiene abierto el surco menor por donde entra la K201, contrario a lo que se ve en un monómero inactivo[13].

Con el fin de poder asociar una curvatura a cada posición de la hélice N, se evaluó el parámetro *roll* del ADN usando el software Curves+[61]. Este parámetro muestra

el grado de rotación de un par de bases respecto al par siguiente, esto lo hace para cada par de bases en la secuencia. Un valor positivo de *roll* lleva a la apertura del ADN en dirección hacia el surco menor, mientras que un valor negativo abre el surco mayor (Figura 15).

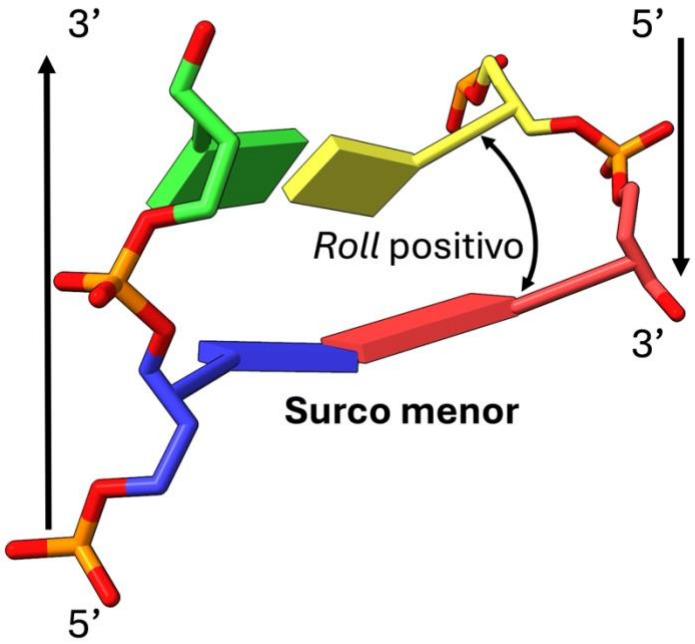


Figura 15. Representación esquemática del parámetro *roll*. Cada base se representa en diferente color, el sentido de la cadena se marca al lado de las bases. El surco menor se muestra de frente dejando al surco mayor en la parte de atrás. La flecha de en medio muestra el movimiento de apertura de un par de bases respecto al otro en el surco menor (*roll* positivo). Figura realizada en ChimeraX 1.9[62].

Se midió el valor de *roll* para la secuencia espaciadora tomando en cuenta regiones que tienen el mismo comportamiento de la hélice N a lo largo del tiempo. Se marcaron los valores de este parámetro para las secuencias *loxP* y *loxI* en el contexto de un intasoma (cuadros magenta y triángulos rojos en la Figura 16, respectivamente); estas estructuras difieren en la dirección de la secuencia espaciadora y en el valor de *roll* en el paso Cyt24-Ade25 (par siete encerrado en

una caja morada en la Figura 16). También se agregaron los valores iniciales de la secuencia *loxI* en el contexto del dímero (círculos negros en Figura 16).

La Figura 16A muestra que el valor de *roll* es simétrico cuando ambas cavidades están ocupadas por la hélice N. Esto sucede durante los microsegundos 1 al 5 de la réplica 1 y en el último microsegundo de la réplica 4; tiempo en el cual los contactos realizados entre el extremo C-terminal y la cavidad adyacente se mantuvieron sin cambios. Particularmente, hubo un incremento en dicho valor en el segundo y en el séptimo par de bases del espaciador (par encerrado en morado, Figura 16A). Por otro lado, cuando la cavidad de Cre^{Yang} está vacía y la cavidad de Cre^{Yin} está ocupada por la hélice N de Cre^{Yang} (intercambio de hélices, Figura 16B), hay un incremento en el valor de *roll* en la posición siete del espaciador (par encerrado en morado), acercándose a los valores vistos en la secuencia tipo intasoma proveniente de 1Q3U (cuadro magenta en el gráfico). El comportamiento observado podría dar indicios sobre el efecto de la posición de la hélice N sobre la curvatura de *loxI*.

Interesantemente, en el caso en el que la hélice N de Cre^{Yin} no salió de la cavidad y la hélice N de Cre^{Yang} no interactuó con la cavidad de Cre^{Yin} (Figura 16C), se observó un comportamiento diferente al esperado, ya que el par siete del espaciador (par encerrado en morado) llega a valores similares a los vistos en el panel 16B, sin embargo, el resto de las bases se mantienen cerca de los valores iniciales (puntos negros), esto podría explicarse con la disminución del número de contactos que hace la hélice N de Cre^{Yin} con la cavidad vecina que, si bien no salió de la cavidad, el número de contactos es menor comparado con las otras réplicas, por otro lado, la hélice N de Cre^{Yang} podría estar interactuando con otras regiones de la proteína o del ADN promoviendo este cambio en los valores de *roll*.

Lo visto en la figura 16A y 16B muestran que, cuando la hélice N de Cre^{Yang} (verde) está invadiendo a Cre^{Yin} (azul), el valor de roll es positivo en la posición 7 del espaciador acercándose a una conformación tipo 1Q3U, mientras que las bases en

la posición 2 parecen no tener cambios en los valores de *roll* a pesar del cambio de posición de la hélice N.

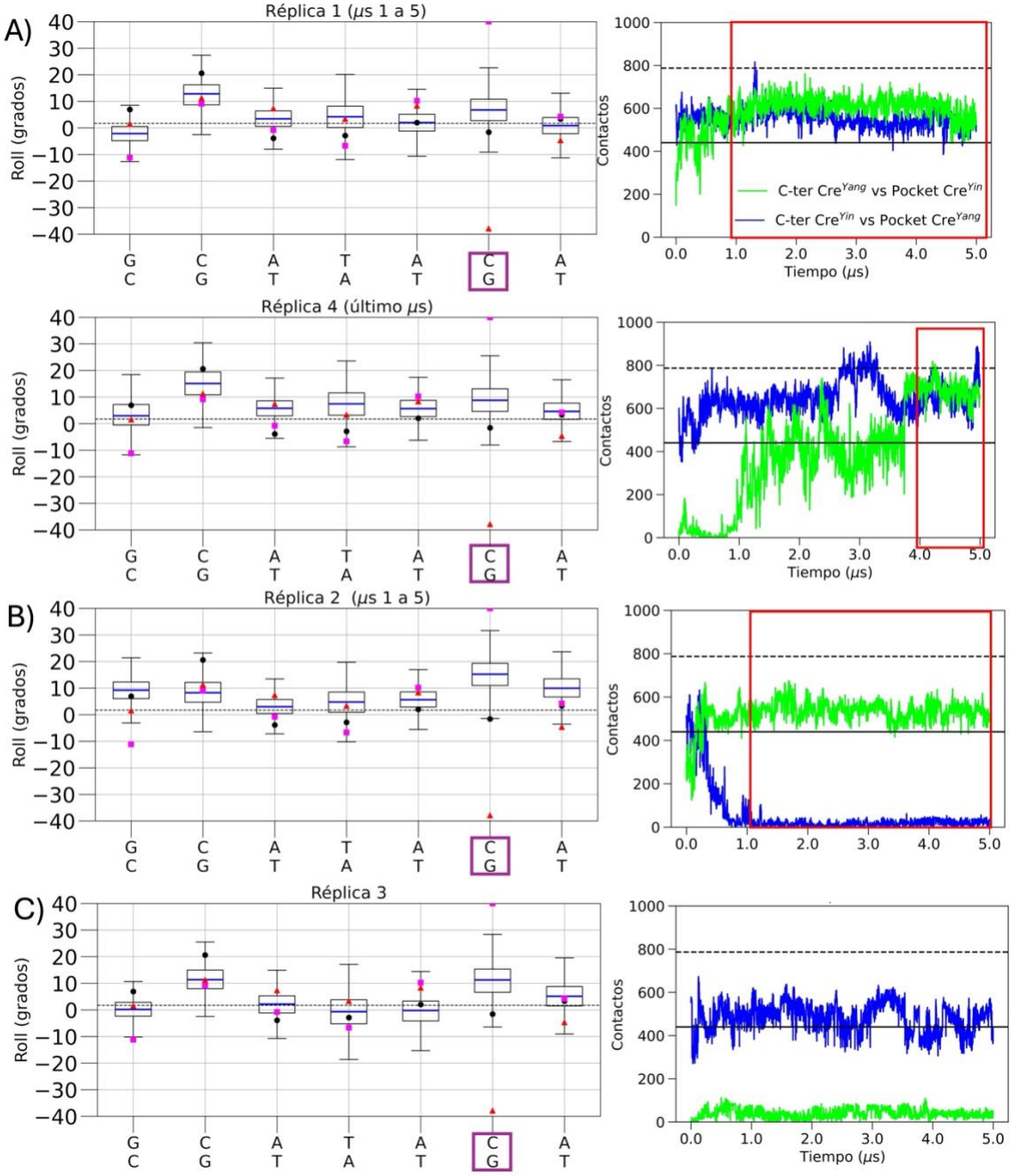


Figura 16. Valor de *roll* asociado a la posición de la hélice N. Gráficas de caja que muestran la A) comparación de *roll* cuando ambas cavidades están ocupadas por

la hélice N (caso 2 Figura 14B), B) cuando hay intercambio de hélice (caso 3, Figura 14B) y C) cuando no hay intercambio de hélice (caso 1, Figura 14B). En todos los casos los valores iniciales de *roll* se muestran en círculos negros (*loxI* dímero), cuadros magenta (*loxP* intasoma) y triángulos rojos (*loxI* intasoma). A la derecha de cada gráfica se muestra el trazo de contactos entre la hélice N y la cavidad adyacente (Cre^{Yin} : azul; Cre^{Yang}: verde).

Por otro lado, la réplica 4 (Figura 17), que tiene un comportamiento diferente a lo largo de diversos intervalos, no muestra una tendencia marcada en los cambios del valor de *roll* en las regiones analizadas. Esto podría indicar que *loxI*, además de cambiar el valor de *roll* en la posición 7 siguiendo una tendencia impuesta por la posición de la hélice N, tiene flexibilidad intrínseca que le permite cambiar de conformaciones sin importar de qué lado esté interactuando cualquiera de los segmentos C-terminales de Cre.

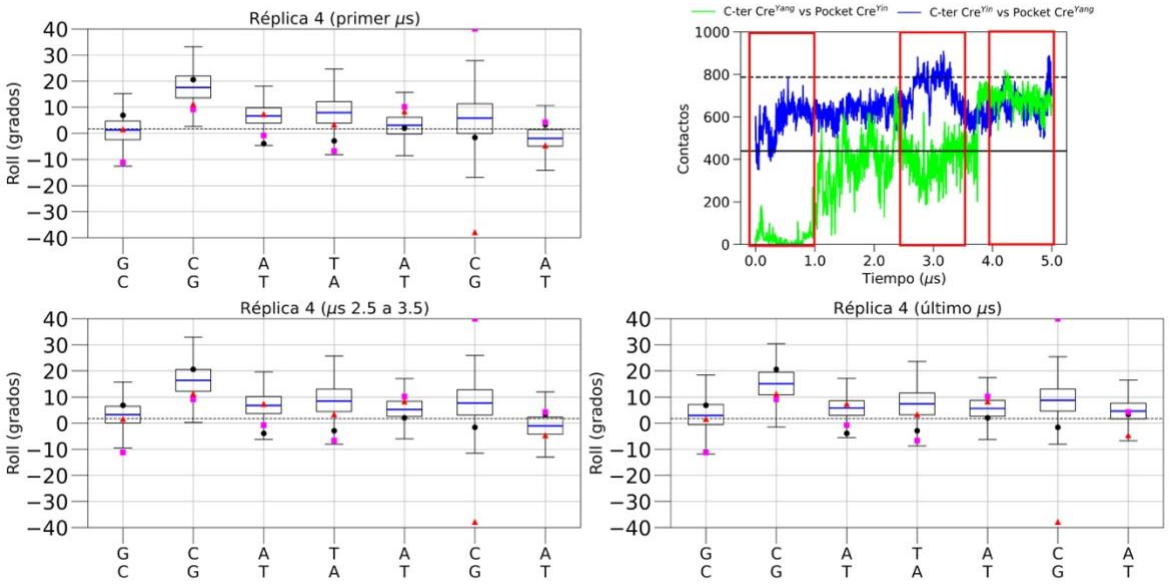


Figura 17. Valor de *roll* de *loxI* independiente a la hélice N. El panel superior derecho muestra los contactos entre el extremo C-terminal de cada monómero en diferentes intervalos a lo largo la simulación. La distribución de valores de *roll* se muestran como gráficas de cajas. En rojo se marcan las 3 regiones analizadas:

primer microsegundo (panel superior izquierdo), región entre 2.5 y 3.5 microsegundos (panel inferior izquierdo), y último microsegundo (panel inferior derecho).

Para tener un mejor panorama del doblez asociado a la posición de la hélice N en la posición 7 del espaciador, la Figura 18 muestra el alineamiento de las estructuras iniciales tomando como referencia uno de los RBEs. La estructura con ambas hélices ocupando la cavidad adyacente también se muestra, ya que correlaciona con el intermediario propuesto por Stachowski et at, en el cual la dirección del doblez puede cambiar según la posición de la hélice N[14].

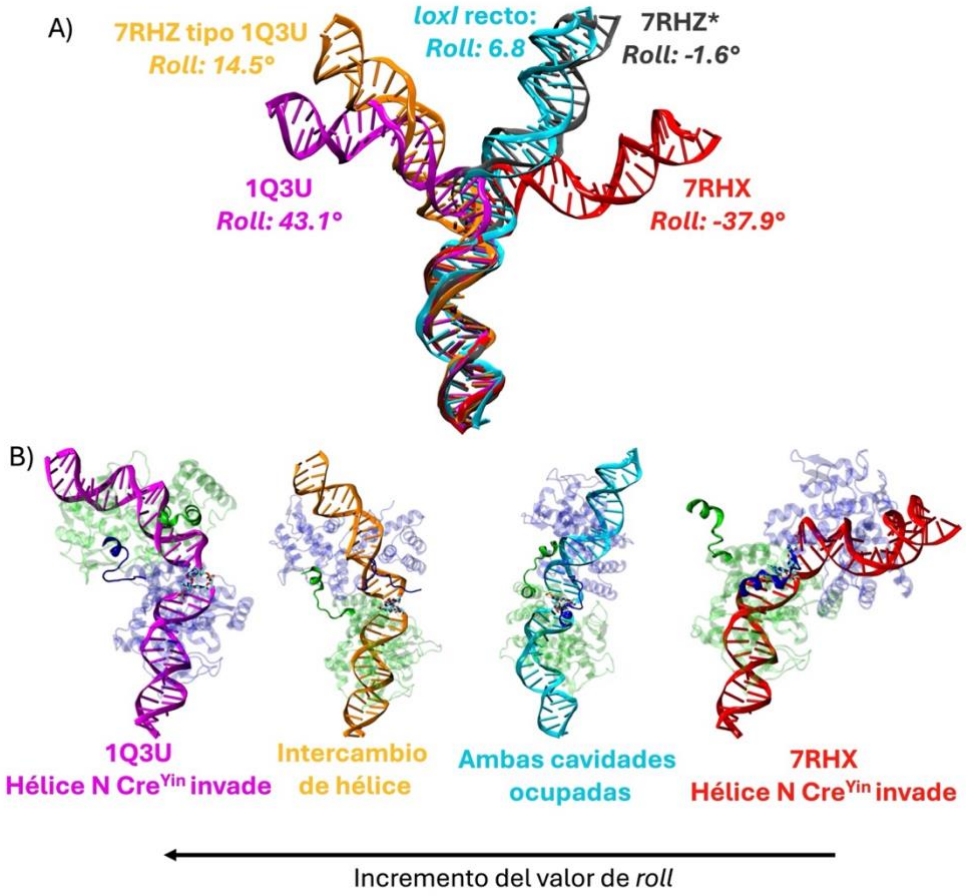


Figura 18. Doblez de *loxI* por cambio en el valor de *roll.* A) Superposición del RBE derecho de las estructuras provenientes de 1Q3U (magenta) con la estructura

7RHX (rojo), *loxI* inicial (gris) y dos intermediarios encontrados en la simulación: loxI recto (ambas hélices ocupado la cavidad adyacente; azul) y loxI en conformación tipo 1Q3U (*roll* positivo, anaranjado). B) Estructuras separadas de la superposición en el panel A mostrando la posición de la hélice A en cada caso. Para cada estructura se muestra el valor de *roll.*

7.1.4 Interruptores moleculares

Actualmente se cuenta con estructuras experimentales de Cre en distintos pasos de la reacción de recombinación, en su mayoría, complejos tetraméricos e intermediarios de Holliday. Por otro lado, también se cuenta con dos complejos triméricos unidos a ADN (1F44, 1DRG), un dímero unido a *loxI* (7RHZ), y un monómero unido a medio sitio *loxA* (7RHY). En el monómero, dímero y trímeros, la posición de la hélice A en el dominio CBD tiene una inclinación de ~60° en dirección al ADN, la cual difiere a la posición vista en el intasoma. Esta hélice interactúa con las hélices C y E del monómero adyacente; mientras que, en el dímero estas interacciones se pierden debido al cambio de posición. Dicho comportamiento lleva a suponer que este segmento podría funcionar como un interruptor molecular durante los últimos pasos del armado del intasoma.

Para evaluar si este cambio de posición ocurre en un complejo dimérico, se calculó la distancia y el ángulo formando entre las hélices A y B, así como la estructura secundaria para este mismo segmento. La Figura 19 muestra que, la hélice A muestrea cuencas cercanas a la conformación inicial (141°, 10 Å y 130°, 10.6 Å, valores de ángulo y distancia para Cre^{Yin} y Cre^{Yang}, respectivamente) pero también se desliza sobre el resto del dominio CBD. Además, la hélice A también puede perder y ganar estructura secundaria (Figura 19C y 19D), lo que hace a este segmento un elemento dinámico en la estructura.

A pesar de poblar diferentes cuencas, la hélice A del monómero Cre^{Yang} sí visita ángulos correspondientes a los del intasoma (-149°, 13.2 Å, 154°, 13.5 Å, valores de ángulo y distancia para un monómero activo e inactivo, respectivamente), sin

embargo, la distancia entre hélices sigue siendo corta. Por otro lado, la hélice A del monómero Cre^{Yin} no es capaz de llegar a estructuras tipo intasoma, ya que se mantiene en posiciones cercanas a las iniciales. Es importante mencionar que este segmento en Cre^{Yin} no cuenta con un monómero vecino, a diferencia del monómero Cre^{Yang}. Esto sugiere que son necesarias interacciones con el monómero adyacente para cambiar de conformación. Otro aspecto que tomar en cuenta es que la hélice A tiene residuos positivos que pudieran estar interactuando con los fosfatos del ADN, dificultando el cambio de conformación.

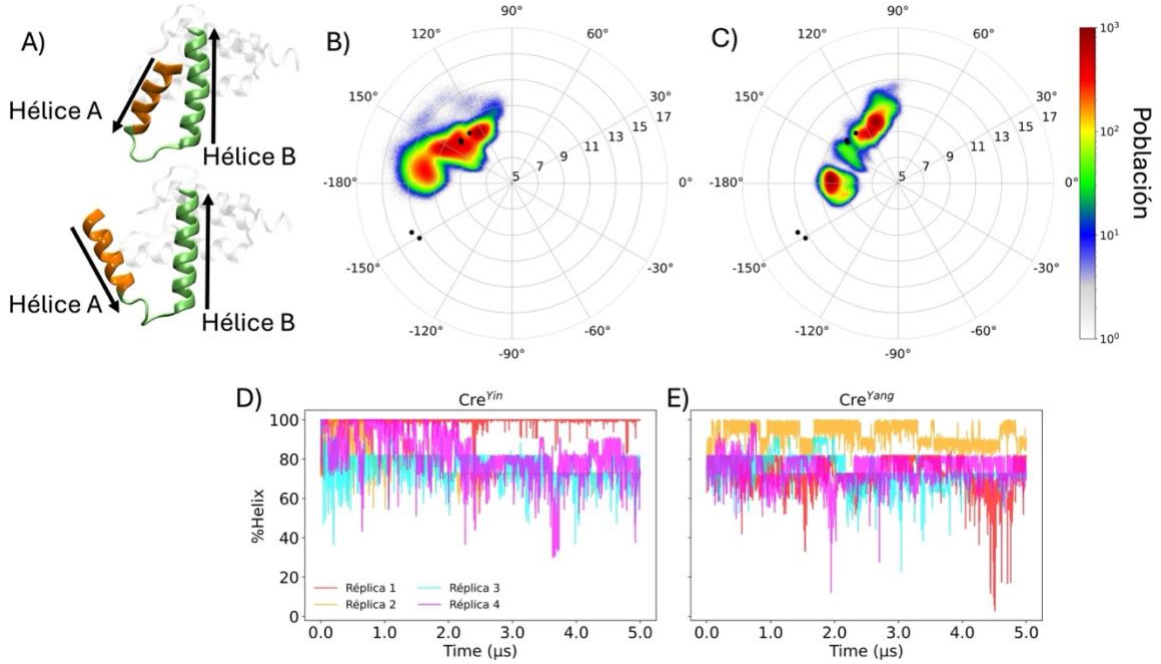


Figura 19. Desplazamiento de la hélice A. A) Definición del ángulo entre las hélices A (anaranjado) y B (verde) para una conformación tipo monómero/dinero (panel superior) y un intasoma (panel inferior). Las flechas indican la dirección del vector usado para el cálculo en sentido N- a C- terminal. B) Paisaje conformacional de la hélice A para Cre^{Yin} y B) Cre^{Yang}, construido con los valores de ángulo y distancia obtenidos. La distancia en Å se muestra del centro hacía afuera del círculo, los ángulos se muestran alrededor del círculo. Los valores iniciales (7RHY, 7RHZ y 1Q3U) se muestran como puntos negros dentro del gráfico. La escala de color muestra el número de miembros en cada bin. C) Porcentaje de estructura secundaria a lo largo del tiempo para Cre^{Yin} y D) Cre^{Yang}.

Para evaluar la disponibilidad de la cavidad en la que interactúa la hélice N se calculó la distancia y el ángulo formado entre las hélices I y K. Los resultados muestran que estas hélices se mantienen cerca de los valores iniciales (Figura 20, puntos negros en el histograma 2D) pero puede abrir y cerrar dicha cavidad. Este movimiento podría facilitar o impedir la interacción con la hélice N.

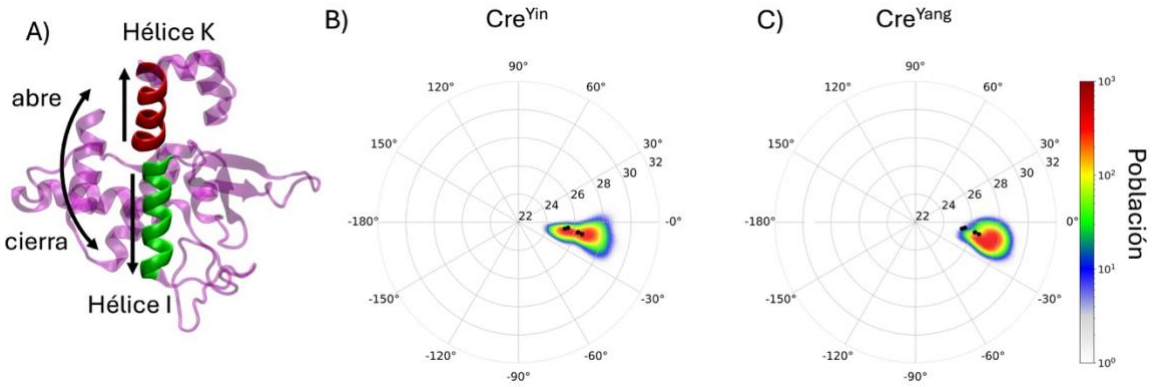


Figura 20. Paisaje conformacional de las hélices I-K. A) Representación de la bisagra en el dominio CAT formada por las hélices I (verde) y K (rojo). Las flechas indican la dirección del vector usado para definir el ángulo. La fleca curva indica el movimiento de apertura y cierre visto para esta bisagra. Paisaje conformacional construido con los valores de ángulo y distancia de las hélices descrita para B) Cre^{Yin} y C) Cre^{Yang}. Los valores iniciales (7RHY, 7RHZ y 1Q3U) se muestran como puntos negros dentro del histograma. La distancia (Å) entre hélices se marca desde el centro hacia afuera del círculo, mientras que los valores de los ángulos se muestran en sentido contrario a las manecillas del reloj. La escala de color muestra el número de miembros en cada bin.

En particular, Cre^{Yin} muestra una distribución de distancias cercanas a estructuras tipo monómero, dímero e intasoma (Figura 20A) además de valores pequeños, lo que sugiere que esta cavidad puede cerrarse al no interactuar con la hélice N. Por otro lado, Cre^{Yang} muestra valores más grandes de distancia y ángulos, esto puede deberse a la interacción de la cavidad con la hélice N lo cual podría promover un incremento en flexibilidad.

El movimiento tipo bisagra identificado (Figura 20A) en el análisis previo muestra que las hélices K, L y M se mueven como un bloque independiente. La hélice M contiene a la tirosina catalítica, por lo tanto, cambios en la posición de este subdominio podrían tener un efecto sobre la formación de un sitio activo funcional.

El movimiento de este subdominio, así como la interacción de la hélice N con la cavidad formada podría tener un efecto sobre la estabilidad de esta tirosina catalítica, sugiriendo que múltiples interacciones son necesarias para disminuir la flexibilidad intrínseca de la proteína para asegurar que todos los residuos que participan en la formación del sitio activo estén en el lugar adecuado.

7.1.5 Integridad del sitio activo

Para que la reacción de recombinación pueda llevarse a cabo, es necesario que los residuos del sitio activo estén en la posición adecuada sobre el fosfato escindible. Estudios previos han mostrado que esta reacción sólo ocurre cuando está formado el complejo tetramérico[60,12].

Como se mencionó en la sección anterior, el residuo catalítico Y324 se encuentra en el subdominio KLM y el movimiento de este subdominio puede alejarlo del resto de los residuos de la constelación catalítica. Por otro lado, la K201, otro residuo clave, se encuentra en la horquilla β2/3, región que ha sido descrita como intrínsecamente desordenada lo que dificulta su reclutamiento hacia el fosfato escindible debido a la pérdida entrópica que esto significa.

Para evaluar la formación de un sitio activo funcional se calculó el radio de giro para los átomos de las cadenas laterales de los residuos del sitio activo que interactúan con el fosfato escindible (R173:NE, NH1, NH2; K201: NZ; H289: ND1, R292: NE, NH1, NH2; W315: NE1, Y324: OH) y los valores obtenidos se compararon con los valores de radio de giro para un sitio activo proveniente de un monómero activo y uno inactivo (cadenas F y A de 1Q3U, respectivamente). La distribución está sesgada hacia valores altos de radio de giro (Figura 21), lo que demuestra que ensamblar el sitio activo no es un evento frecuente, además, solo Cre^{Yin} muestrea

valores similares a los de un monómero activo (línea punteada 4.17Å en Figura 21A). Sin embargo, a pesar de esto la posición de las cadenas laterales no corresponde a la de un sitio activo funcional como se muestra en las Figuras 21B y 21C, lo que sugiere que se necesita un ajuste extra de las cadenas laterales. Lo mismo sucede con el monómero Cre^{Yang}, el cual no logra muestrear valores similares a un monómero activo, pero sí a los de uno inactivo. Nuevamente, a pesar de ser valores similares, la orientación de cadenas lateral difiere a la de la estructura cristalográfica. Esto confirma que no solo es necesario reclutar a todos los residuos del sitio activo alrededor del fosfato escindible, sino que también hay que ajustarlos en la posición correcta para que los grupos funcionales relevantes estén en la orientación adecuada para llevar a cabo el ataque nucleofílico y estabilización de los intermediarios de la reacción.

Como se mencionó anteriormente en la sección 7.1.1. la posición de la horquilla β2/3 en la que reside la K201 está alejada de *loxI* y del resto de los residuos del sitio activo. Sin embargo, a lo largo del tiempo de simulación, esta horquilla es capaz de ser reclutada hacia posiciones que permiten tener valores de radio de giro similares a los de un monómero activo e inactivo. Este muestreo también indica que el tiempo de simulación es suficiente para observar múltiples conformaciones de las regiones desordenadas de Cre.

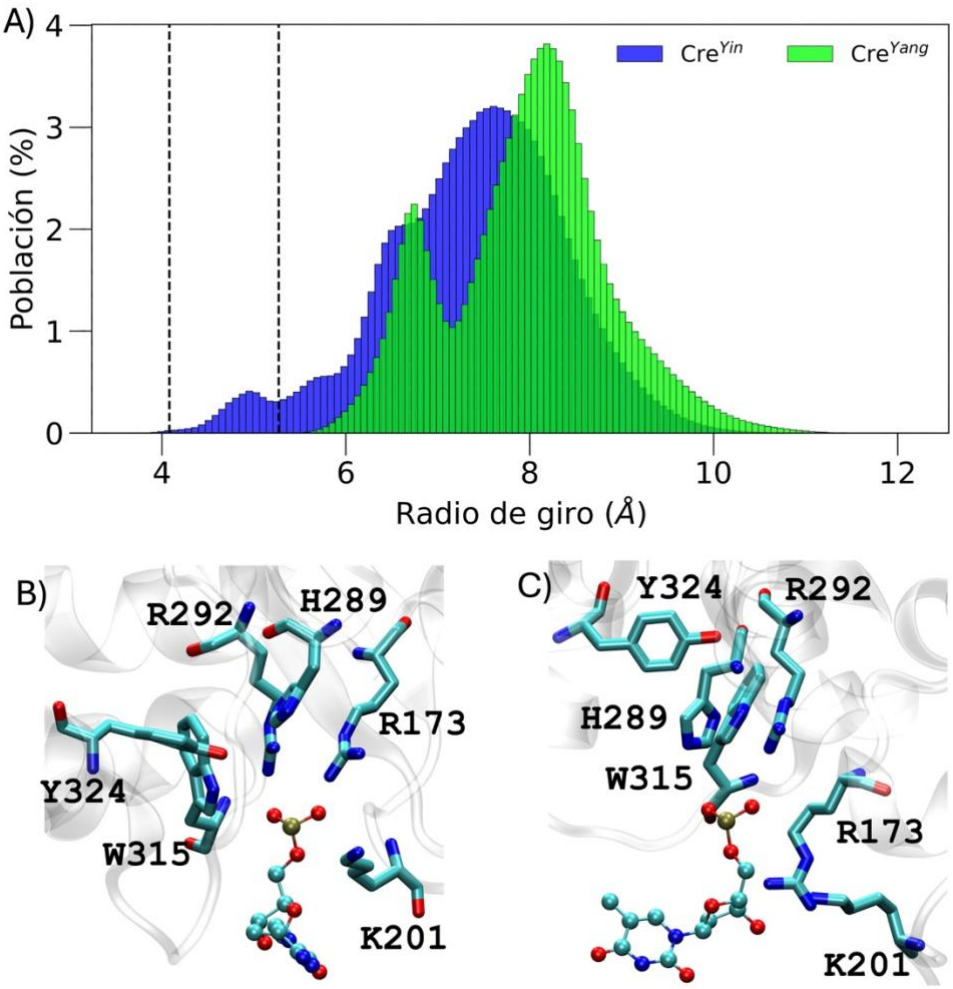


Figura 21. Ensamble del sitio activo. A) Histograma de radio de giro de los átomos de las cadenas laterales del sitio activo alrededor del fosfato escindible para cada monómero; las líneas punteadas representan los valores de radio de giro del sitio activo para un monómero activo (4.17 Å) y uno inactivo (5.22 Å). B) Estructura con el radio de giro más pequeño para Cre^{Yin} (3.65 Å) y C) Cre^{Yang} (5.26 Å), los residuos del sitio activo se etiquetan en cada panel. Las estructuras están alineadas respecto al fosfato para mostrar las diferentes posiciones de las cadenas laterales de la constelación catalítica.

Por otro lado, para observar con mayor detalle las interacciones realizadas entre el sitio activo y el fosfato escindible, se calcularon estas interacciones con el software LigPlus[63], este programa nos permite observar cuales son las interacciones tipo puente de hidrógeno, electrostáticas e hidrofóbicas de una proteína con su ligando. La Figura 22 muestra que a pesar de visitar valores pequeños de radio de giro el sitio activo no se ensambla como en un monómero activo, lo que sugiere que en el contexto del dímero, Cre no forma un sitio activo funcional y, por ende, no lleva a cabo la reacción de recombinación.

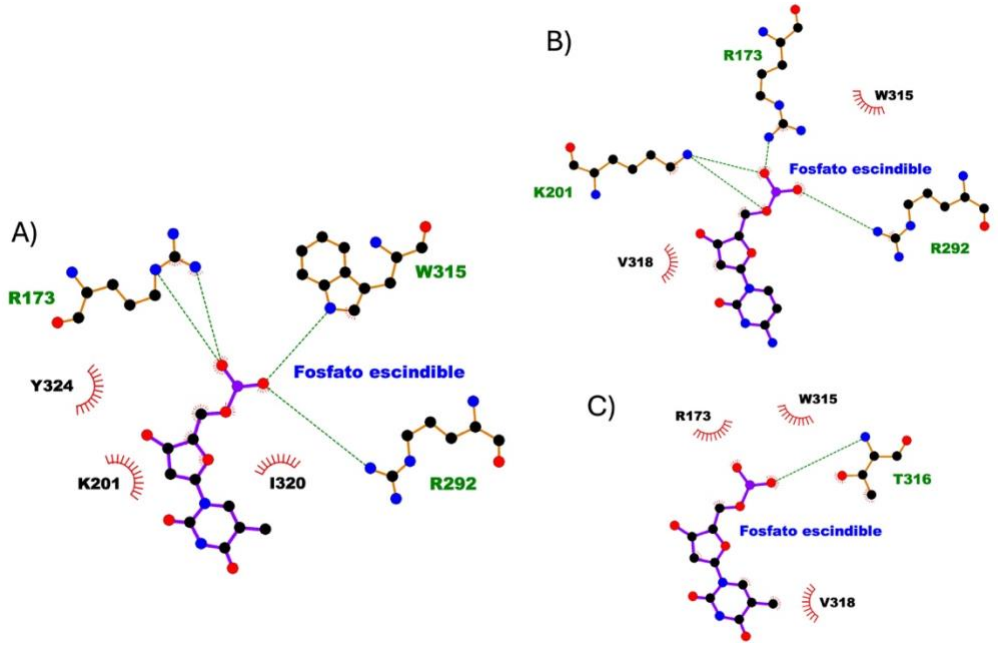


Figura 22. Detalle de las interacciones con el sitio activo. A) Interacciones realizadas en el monómero activo de Cre (PDBID 1Q3U:F). B) Interacciones de los residuos del sitio activo con la estructura con el radio de giro más pequeño visitado en Cre^{Yin} y C) Cre^{Yang}.

7.2 Cre libre y dominios aislados

7.2.1 Validación experimental

Para evaluar si el ensamble generado describe adecuadamente la dinámica de Cre libre, se calcularon los desplazamientos químicos con el software UCBShift[64] (CS, por sus siglas en inglés) y se compararon con datos experimentales. Hasta la fecha, solo se cuentan con CS para el dominio CAT disponibles para la comparación.

Los resultados muestran que hay una buena correlación entre los datos predichos y los experimentales para los átomos Cα y Cβ, con valores de R^2 de 0.948 y 0.995, respectivamente (Figura 23).

Estos resultados muestran que el tiempo de simulación es suficiente para poder recrear resultados experimentales. Esto nos permite proponer explicaciones que ayuden a entender procesos que no han sido del todo explicados, así como proponer nuevos experimentos que respondan a nuevas preguntas hechas para este sistema.

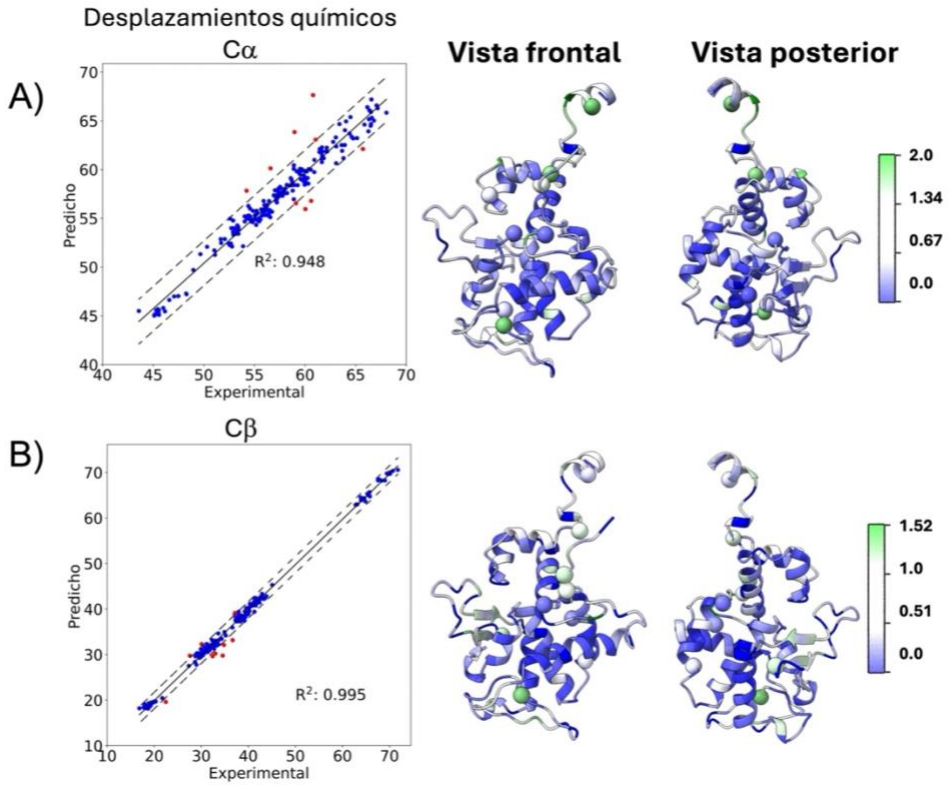


Figura 23. Comparación con datos experimentales: desplazamientos químicos de Cα y Cβ. Se muestra la comparación pareada de los desplazamientos químicos calculados con los experimentales para A) Cα y B) Cβ. En rojo se marcan aquellos datos que se salen del 95% del intervalo de confianza (dos desviaciones estándar). A la derecha se muestra la estructura del dominio CAT proveniente de Cre completa, pintada de azul a verde según los valores de desviación estándar. Los puntos rojos de cada gráfica se marcan como esferas en las estructuras.

7.2.2 Potencial electrostático y radio de giro

Para evaluar el comportamiento de Cre en ausencia de interacciones proteína-proteína y proteína-ADN se calculó el radio de giro tomando solamente los Cα. Los valores obtenidos se compararon con los valores de radio de giro de un monómero activo e inactivo, los dos monómeros de 7RHZ y el monómero de 7RHY (Figura 24B). Además, analizamos el potencial electrostático con APBS[65] a pH 7 y las cargas de CHARMM (Figura 24A) con el fin de encontrar regiones cargadas que estuvieran expuestas. Estas regiones podrían contribuir a un incremento de flexibilidad a causa de repulsión o atracción electrostática.

Los resultados muestran que Cre tiene una amplia distribución de valores de radio de giro, desde ~21Å (más compacta que en un intasoma) hasta ~40Å. Además, el potencial electrostático muestra que Cre expone dos superficies cargadas positivamente (superficies azules en la estructura), las cuales corresponden a las regiones de los dominios CBD y CAT que están en contacto con el ADN. Adicionalmente, el potencial electrostático de Cre de cada uno de los picos mostrados en el histograma (Figura 24C), muestra que, efectivamente, los parches positivos se repelen entre sí. Sin embargo, en la superficie también hay parches negativos (superficies rojas) que podrían promover una interacción dada por complementariedad de cargas.

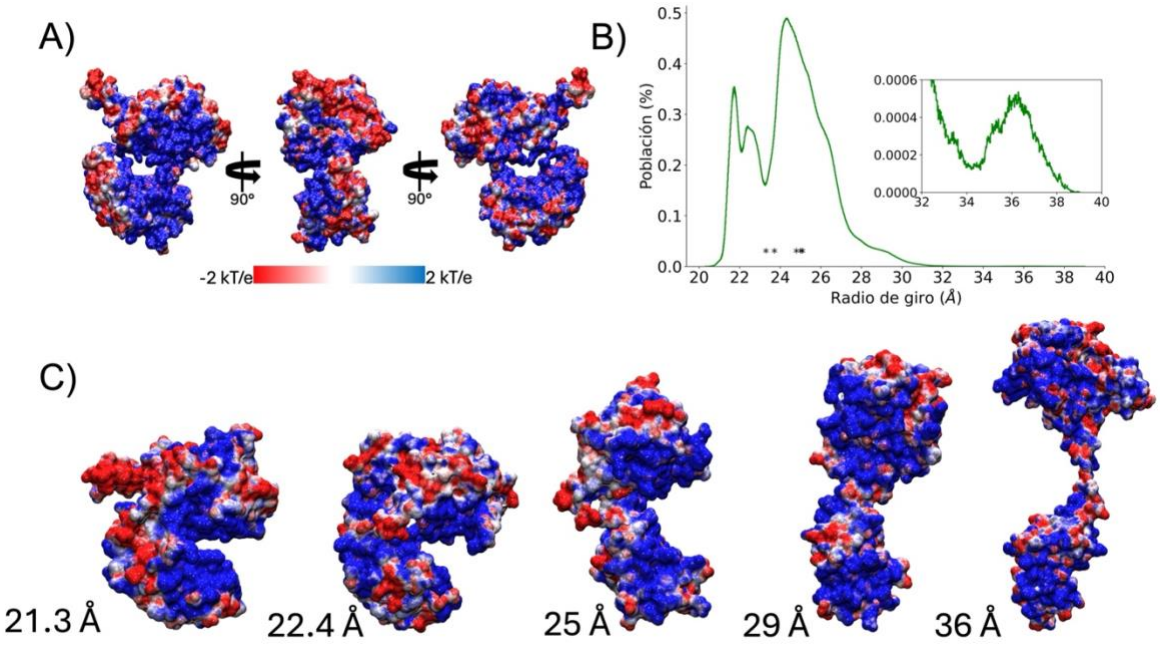


Figura 24. Potencial electrostático y distribución de radio de giro. A) Potencial electrostático de Cre en ausencia de interacciones proteína-proteína y proteína-ADN mostrada en tres posiciones diferentes. B) Distribución de radio de giro para todo el ensamble; los asteriscos muestran los valores de las cinco estructuras iniciales, mientras que en el acercamiento muestra los valores extremos obtenidos. C) Potencial electrostático marcado sobre la superficie de las estructuras correspondientes a cada pico del histograma del panel B. Los valores de radio de giro para cada estructura se muestran debajo de cada estructura.

7.2.3 Flexibilidad del conector entre dominios

Los dominios CBD y CAT están unidos entre sí por un conector flexible, el cual ha mostrado ser importante para el correcto acomodo de Cre sobre el ADN ya que permite el movimiento de un dominio respecto al otro. Sin embargo, este conector también es estabilizado por interacciones con el ADN. Para estudiar la flexibilidad

de esta región se calcularon los ángulos de cadena principal (phi y psi) para construir un mapa de Ramachandran.

Los resultados muestran que, como era de esperarse, el conector visita todas las regiones permitidas en el mapa (Figura 25A). Esto correlaciona con lo visto en las estructuras de los picos del radio de giro, en las cuales el conector se encuentra en múltiples conformaciones. Además de esto, se calculó la distancia entre los extremos del conector (Figura 25B), esto con el fin de averiguar si esta distancia tiene un efecto sobre la separación o acercamiento entre los dominios. Al igual que el radio de giro, la distribución de distancias tiene una distribución amplia, con dos poblaciones principales en valores ~20Å y ~27Å. Estos valores son distintos a los de las estructuras iniciales (asteriscos negros en Figura 24B), lo cual muestra que la posición inicial de este conector no es una de las más frecuentes, pero sí es visitada.

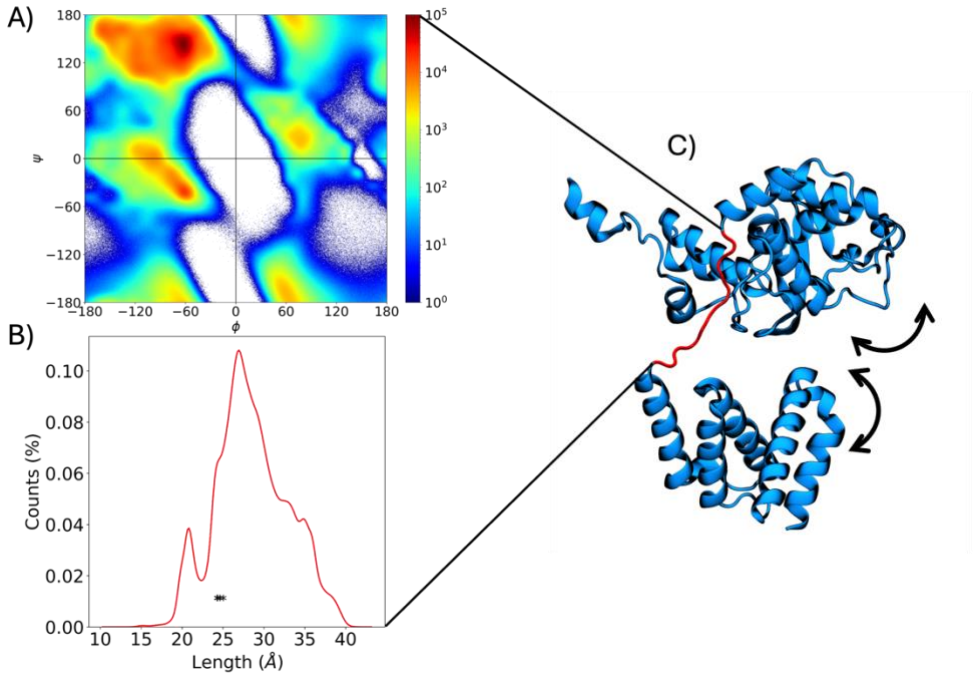


Figura 25. Mapa de Ramachandran y distribución de distancia entre extremos del conector. A) Mapa de Ramachandran para los residuos del conector, la escala de colores muestra el número de miembros en cada combinación de valores. B) Distribución de distancia entre las puntas del conector. Los valores iniciales se

muestran como asteriscos negros sobre el histograma. C) Estructura de Cre resaltando la posición del conector (trazo rojo). Las flechas entre dominios muestran los posibles cambios en orientación de los dominios.

7.2.4 Múltiples orientaciones entre dominios

El radio de giro y la distribución de la distancia entre extremos ha mostrado que Cre puede visitar diferentes estados, ya sea extendidos (valores altos de radio de giro) o colapsados (valores bajos de radio de giro). Sin embargo, estos descriptores no dan información sobre la orientación relativa de un dominio respecto a otro. Para tener acceso a esta información se seleccionó una hélice de cada dominio (la más cercana al centro de masa de cada uno, Figura 26A) y se calculó la distancia en los Cα centrales de cada hélices y ángulo diedro formado entre ellas. Los resultados se compararon con los de un dímero de Cre unido a *loxI* (Figura 26B, 26C, 26D).

La distribución de ángulos y distancias de Cre en ausencia de ADN es amplia, ya que es capaz de visitar la mayoría de los ángulos disponibles, además, la distancia entre estas dos hélices también correlaciona con la separación de los dominios llevando a valores altos de radio de giro. A pesar de muestrear múltiples estados, Cre libre también pasa por aquellas combinaciones correspondientes a las vistas en un monómero, dímero e intasoma (asteriscos rojos en la Figura 26B). Además, cuando se mapean los valores de ángulo y distancia para las estructuras representativas del ensamble (asteriscos negros en la Figura 26B), se observa que hay conformaciones más accesibles que otras, y que estas conformaciones corresponden a las cuencas más pobladas en el paisaje conformacional.

Cuando se comparan estos resultados con los monómeros de Cre^{Yin} y Cre^{Yang} se puede observar que este movimiento disminuye, evitando que los dominios giren uno respecto al otro; demostrando que Cre, en presencia de ADN, se encuentra en una conformación restringida que le impide moverse libremente.

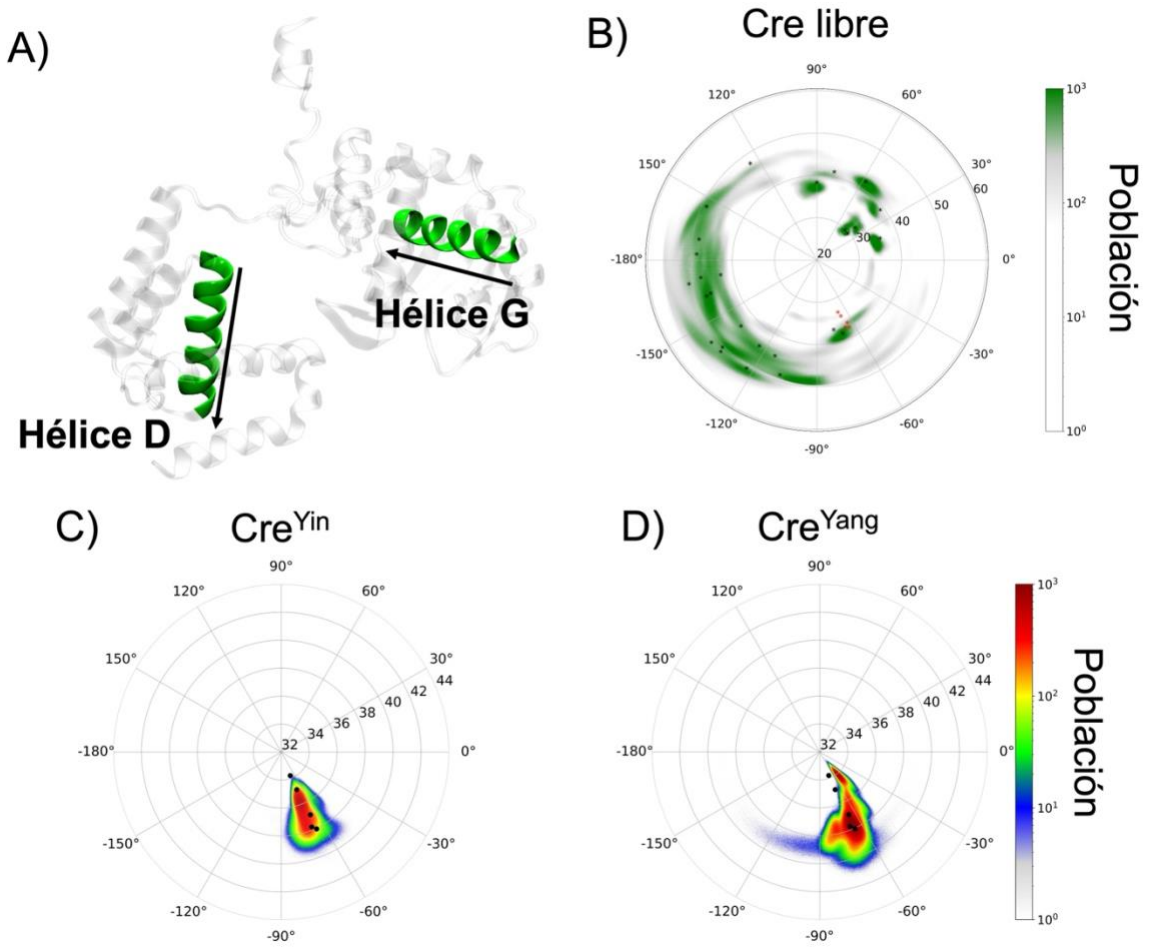


Figura 26. Paisaje conformacional de Cre libre y con *loxI*. A) Estructura de Cre mostrando las hélices elegidas para el cálculo de ángulos y distancias, la dirección de la flecha indica la dirección del vector definido para el cálculo del ángulo diedro (N- ter a C-ter). B) Paisaje conformacional de Cre en ausencia de ADN, los valores iniciales se muestran como asteriscos rojos; mientras que, los valores correspondientes a las estructuras representativas del ensamble se muestran como asteriscos negros (las estructuras de los 5 clústers más poblados se muestran en el Anexo III). C) Paisaje conformacional de Cre^{Yin} y D) Cre^{Yang}, los valores iniciales se muestran como puntos negros sobre cada paisaje. La escala de color muestra el número de miembros en cada bin.

7.2.5 Accesibilidad de superficies de interacción

El movimiento exhibido por Cre en ausencia de *loxI* podría exponer u ocultar las superficies de interacción con el ADN o en las interfaces proteína-proteína. Para evaluar esta accesibilidad, se calculó el área accesible al solvente de cada una de las regiones de interacción proteína-ADN (zonas de CBD y CAT en interacción directa con el ADN, Figura 27B) y proteína-proteína de cada dominio (Figura 27A), definidas de la siguiente manera:

Cara1-CBD: hélice A y hélice B

Cara2-CBD: hélices D, E y una pequeña región de la hélice C

Cara1-CAT: cavidad de interacción con la hélice N

Cara2-CAT: hélice N y extremo C-terminal del conector entre ambos dominios.

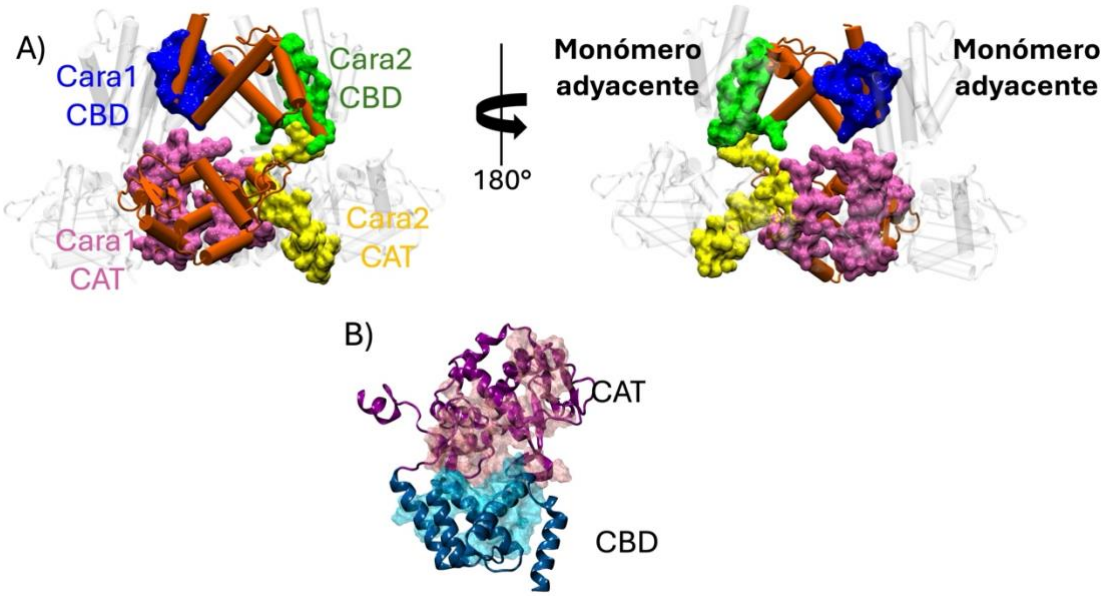


Figura 27. Superficies de interacción proteína-proteína y proteína-DNA. A) regiones de interacción proteína-proteína. Azul: Cara1-CBD; verde: Cara2-CBD, rosa: Cara1-CAT; amarillo: Cara2-CAT. Las subunidades vecinas de Cre se muestran como tubos transparentes. B) Zonas de interacción con el DNA para CBD (cian) y CAT (morado).

El área de estas cuatro regiones de interacción proteína-proteína fue evaluada a lo largo del ensamble con una resolución de 1ns. Las gráficas de violín presentadas en la Figura 28, muestran que la Cara1-CBD tiene una distribución simétrica alrededor del valor inicial. Por otro lado, el área de la Cara2-CBD muestra una distribución que se mantiene cerca del valor inicial pero también muestrea valores menores a la inicial (Figura 28A). Este resultado sugiere que, a pesar de visitar múltiples estados, el área de interacción del CBD con otra proteína, se mantiene accesible al menos durante la mitad del tiempo de simulación.

En el caso del dominio CAT, la Cara1-CAT tiene una distribución que está centrada abajo del valor inicial, además de que también muestrea valores mayores, esta región, que corresponde a la cavidad en la que la hélice N interactúa en el contexto de un dímero/intasoma es una región más dinámica, cambiando su área ya sea arriba o abajo del promedio. En el caso de la Cara2-CAT, la cual corresponde a la hélice N y el C- terminal del conector flexible, muestra una disminución general en la exposición de área (Figura 28B). Este comportamiento correlaciona con estudios previos que muestran que esta región es más rígida de lo esperado, lo que ha llevado a proponer que este segmento puede formar un complejo auto-inhibitorio[11] en el cual la hélice N interactúa con la región de interacción con el DNA bloqueando a los residuos del sitio activo.

Cuando se analizaron las regiones de interacción con el DNA, para los dominios por separado, el área de unión al ADN en el CBD tiene una distribución amplia, en la cual el promedio se encuentra por debajo del valor inicial. En el caso del dominio CAT, el área parece incrementar comparada con el área inicial (Figura 28C). Esto en parte puede ser explicado debido a que, en el contexto de un monómero unido a *loxP,* el dominio CBD esconde parte del dominio CAT, de modo que cuando los dominios cambian de orientación, el área del dominio CAT queda expuesta y por tanto su valor aumenta. Sin embargo, este incremento en área de interacción con el ADN no necesariamente implica que esta región esté lista para poder unirse al ADN y, por ende, llevar a cabo la reacción de recombinación.

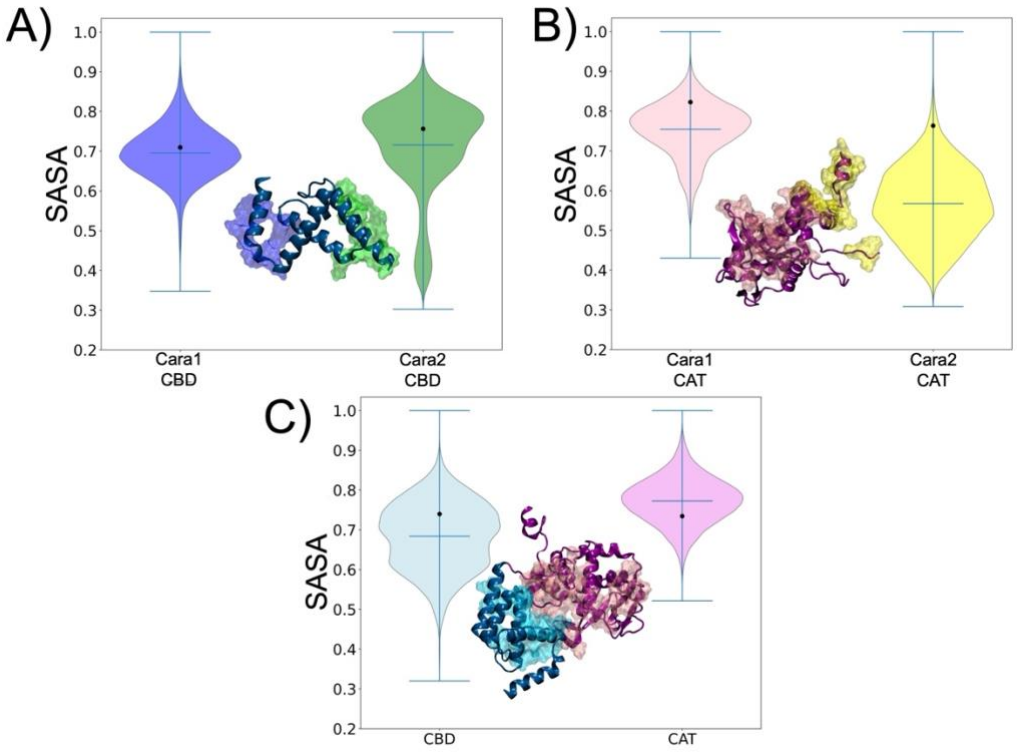


Figura 28. Exposición de áreas de interacción. Gráficas de violín mostrando la distribución de área accesible al solvente (SASA) para A) CBD Cara1-CBD (azul), Cara2-CBD (verde) y B) Cara1-CAT (rosa) y Cara2-CAT (amarillo). C) Distribución de áreas de los residuos correspondientes a la región de interacción con el ADN para CBD (cian) y CAT (morado). En todos los casos, las áreas están normalizadas respecto al valor más alto encontrado en el ensamble. Los valores iniciales se muestran como puntos negros dentro del gráfico. Cada una de las estructuras, junto con el área medida (mostrada como superficie molecular) se muestra dentro de cada gráfico.

7.2.6 Carácter desordenado del extremo C-terminal

Como se mostró en la sección anterior, contrario a lo que se esperaba, el extremo C-terminal (que incluye a la hélice N) muestra una disminución de área accesible al

solvente, una característica que es poco esperable dado el carácter desordenado de este segmento. Para explicar esta diminución en área expuesta y, probablemente en flexibilidad, se calcularon los contactos realizados entre este segmento y la cavidad formada en el dominio CAT, así como con el sitio activo. Este análisis permite caracterizar las interacciones realizadas por el extremo C-terminal que pudieran explicar los resultados mostrados en la Figura 28B. Se definieron dos tipos de cavidades, una usando las interacciones iniciales en un intasoma y otra con las provenientes de un dímero. La distribución de contactos se muestra en la Figura 29. Los contactos realizados con ambas cavidades muestran que el extremo C-terminal tiene cierta preferencia por la cavidad tipo intasoma, ya que la distribución muestra un incremento en valores superiores al promedio. Esto sugiere que la hélice N forma interacciones más estables con la cavidad tipo dímero (Figura 29A, violines azul y verde), además, estas interacciones (dadas en el contexto de un dímero y un intasoma) podrían empezar a dar información sobre un proceso de "*domain swapping*", es decir, el segmento C-terminal puede hacer el mismo tipo de interacciones tanto en su mismo dominio, como con el dominio vecino en un complejo dimérico y/o tetramérico.

También se observan interacciones con el sitio activo (residuos R173, E176, K201, H289, R292, W315, y Y324) (Figura 29A, violín anaranjado). Esto correlaciona con las observaciones experimentales que sugieren la formación de un complejo auto-inhibitorio, sin embargo, este complejo fue caracterizado experimentalmente en ausencia del dominio CBD[11], mientras que en este estudio se usó la proteína completa. El dominio CBD podría tener un efecto sobre la persistencia de estas interacciones. Por otro lado, se calculó la estructura secundaria de este segmento C-terminal. El mapa de Ramachandran (Figura 29B) muestra que el extremo C-terminal muestrea todas las regiones permitidas del mapa, lo que confirma el carácter desordenado de este segmento mostrando que puede adquirir múltiples conformaciones.

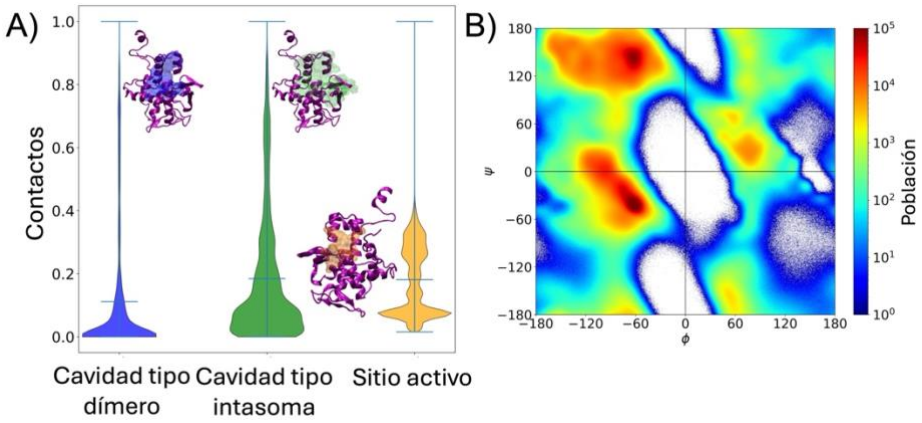


Figura 29. Interacciones del extremo C-terminal y mapa de Ramachandran. A) Distribución de contactos realizados entre el extremo C-terminal y las cavidades tipo dímero, intasoma y los residuos del sitio activo. Los contactos están normalizados contra el número mayor de contactos visto a lo largo de la simulación. B) Mapa de Ramachandran para los residuos del extremo C-terminal.

7.2.7 El conector y el extremo C-terminal bloquean al sitio activo

En la sección previa se mostró que el extremo C-terminal interactúa con las cavidades definidas para un dímero, un intasoma y con los residuos del sitio activo. Sin embargo, al ser este un segmento desordenado también podría interactuar con el resto de la proteína. La inspección de las estructuras obtenidas a lo largo de la dinámica mostró que el conector entre ambos dominios también interactúa con el sitio activo. Se calcularon los contactos hechos por el extremo C-terminal y el conector interdominios con los residuos del sitio activo. Estos valores se compararon con los del sistema CAT, el cual, a diferencia de Cre libre no cuenta con el dominio CBD. Esto tiene el fin de evaluar si la presencia de este dominio tiene un efecto sobre la formación de un complejo auto-inhibitorio.

Los resultados muestran que el extremo C-terminal y el conector flexible interactúan con el sitio activo en ambos sistemas (Figura 30). Sin embargo, a este último, no se le ha atribuido un papel inhibitorio; estas interacciones son más persistentes en

el sistema CAT en comparación con Cre libre (Figura 30B). En el contexto de la formación del intasoma estos segmentos son desplazados por el ADN. Uno de los mecanismos para este desplazamiento podría ser que la gran carga negativa presente en el ADN promueve un cambio conformacional en el dominio CAT desplazando al conector o al extremo C-terminal del sitio de unión. La formación de este complejo auto-inhibitorio es congruente con datos experimentales obtenidos con PRE[11] (Paramagnetic Relaxation Enhancement). Sin embargo, aún no existe una estructura experimental que compruebe este complejo. Por otro lado, nuestras estructuras muestran que el extremo C-terminal no forma estructura tipo hélice cuando está en este complejo, pero al ser una región desordenada cabe la posibilidad de que haya un mecanismo de ajuste inducido una vez formada la interacción.

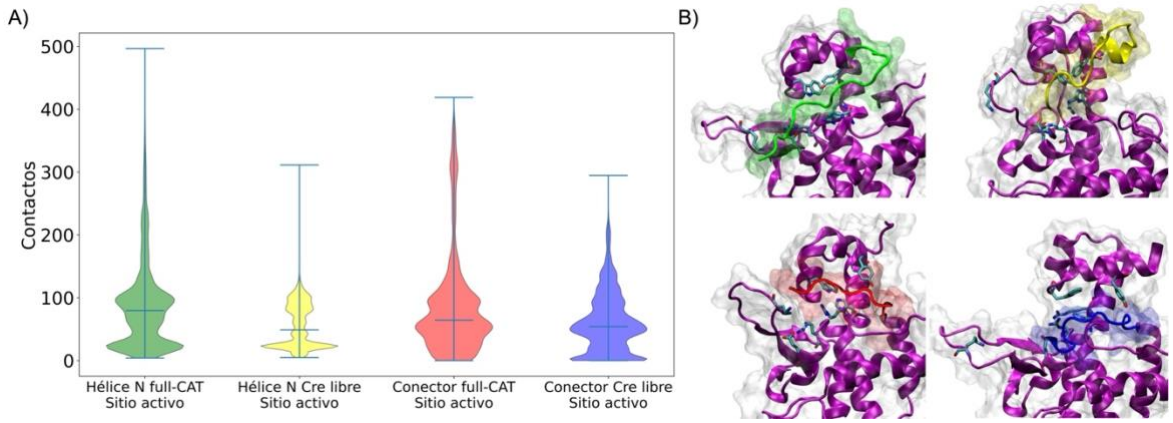


Figura 30. Múltiples interacciones del extremo C-terminal y el conector. A) Gráficas de violín para el extremo C-terminal y el conector con el sitio activo. B) estructuras con el mayor número de contactos realizados entre cada segmento y el sitio activo en ambos sistemas, el color del segmento es el mismo que el de los violines en el panel A.

7.2.8 Integridad del sitio activo en ausencia de *loxI*

Como se mostró anteriormente, el sitio activo puede ser inhibido por la interacción con el extremo C-terminal y el conector interdominios. Sin embargo, un aspecto que aún queda por ser investigado es si este sitio activo puede ensamblarse en

ausencia de ADN. Para este análisis se calculó el radio de giro de los residuos que forman al sitio activo para todos los sistemas sin ADN y se compararon los resultados con los valores de un sitio activo proveniente de un monómero en conformación activa (1Q3U:F). Para Cre libre, también se graficaron los mapas de Janin para los ángulos χ1 y χ2 de cadena lateral de estos residuos y se compararon con las mismas cadenas lateral de los sistemas unidos a *loxI* (Cre^{Yin} y Cre^{Yang}). Esto permite ver qué tan diferente es la posición de las cadenas laterales de estos residuos cuando no hay interacciones proteína-ADN.

La distribución de radio de giro (Figura 31A) de Cre libre es amplia, con valores desde ~6 Å (menor al valor de un sitio activo funcional) y hasta ~16 Å. Esto sugiere que en este sistema el sitio activo se puede sobre comprimir llevando a conformaciones que no correlacionan con un sitio activo funcional. Sin embargo, sí se visitan conformaciones funcionales.

CAT tiene una distribución más compacta, visitando pocos valores por debajo de 6.8 Å hasta un máximo de ~13 Å, A diferencia del sistema CAT-delN, el cual tiene la distribución más amplia, ya que llega a valores de hasta ~21 Å, además de tener un pico alrededor de los 18 Å, valor que ninguno de los otros dos sistemas alcanzaron durante el tiempo de simulación. Este sistema no posee los residuos correspondientes al conector interdominios y el extremo C-terminal. La ausencia de estas regiones podría tener un efecto sobre el aumento de flexibilidad de este sistema. En todos los sistemas, la distribución está sesgada principalmente hacia valores grandes de radio de giro. Este comportamiento está dado por la posición de la K201, residuo que se encuentra en la horquilla β2/3, la cual es una región intrínsecamente desordenada, promoviendo que este residuo se aleje del resto de sitio activo, llevando a valores altos de radio de giro (Figura 31B).

Este análisis mostró que, en ausencia de ADN, la formación de un sitio activo es un evento poco frecuente tanto en presencia (Figura 21) como en ausencia de ADN (Figura 31), nuevamente sugiriendo que son necesarias las interacciones con los demás monómeros. Además, la distribución mostró que la presencia o ausencia del

dominio CBD no tiene un efecto sobre el ensamblado del sitio activo, esto debido a que incluso en las estructuras con radio de giro más compacto, el dominio CBD no está interactuando con los residuos del sitio activo que pudieran explicar esta disminución en radio de giro (Figura 31C). A pesar de que es poco frecuente, sí se visitan valores correspondientes a un sitio activo funcional, sin embargo, este ensamblado parece ser estable solo cuando hay interacciones con otros monómeros y el ADN, es decir, en el contexto de un intasoma.

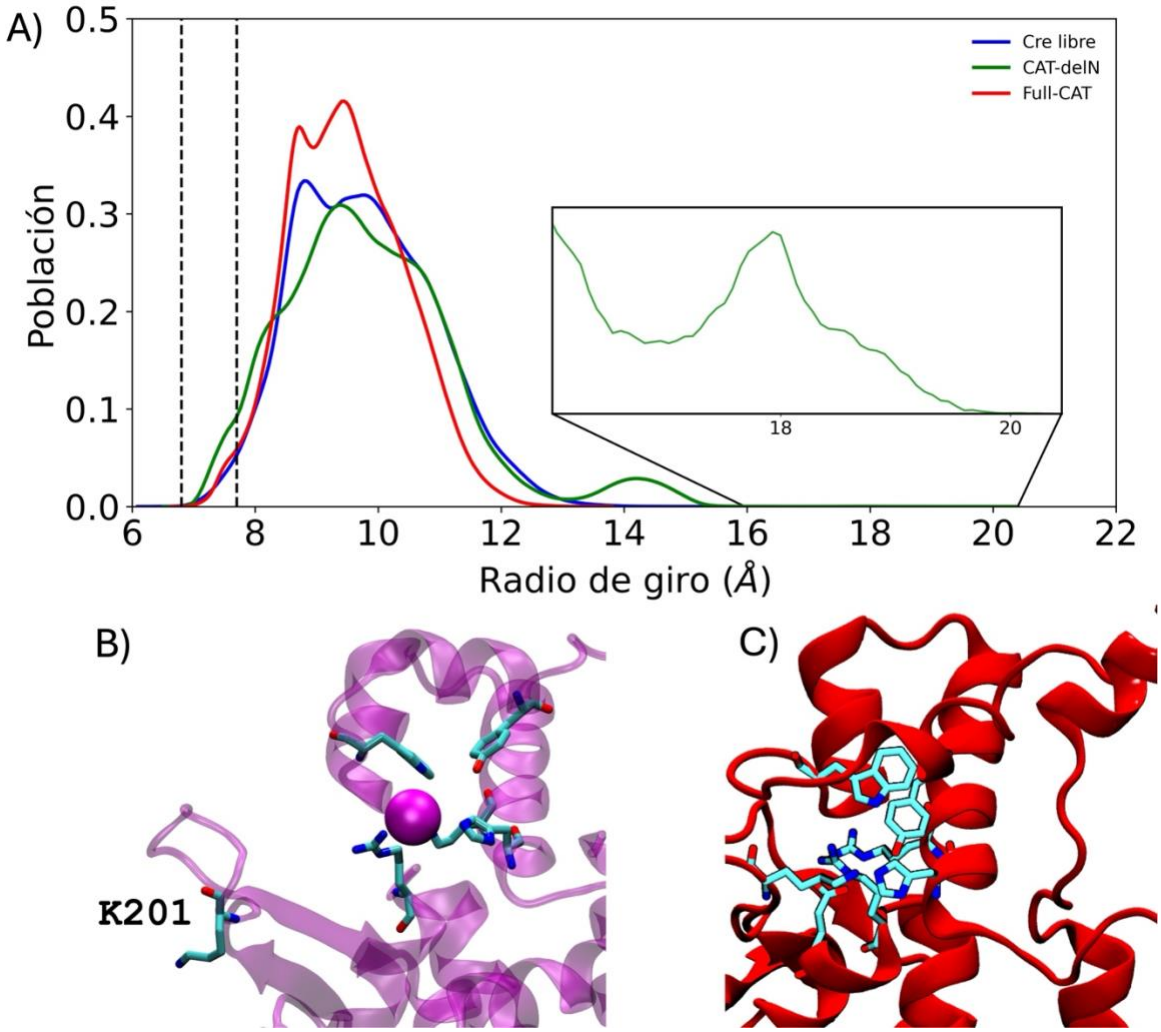


Figura 31. Distribución de radio de giro del sitio activo. A) Histograma de radio de giro para Cre libre (trazo azul), CAT-delN (trazo verde) y CAT (trazo rojo). Los

valores de radio de giro para un monómero activo y uno inactivo se muestran como líneas punteadas en los valores 6.8 Å y 7.7 Å, respectivamente. El acercamiento muestra los valores extremos muestreados por CAT-delN. B) Residuos del sitio activo marcados sobre la estructura del dominio CAT; la esfera representa el centro de masa de estos residuos. La K201 se etiqueta y se muestra alejada de este centro de masa. C) Estructura de Cre libre con el radio de giro más pequeño (6.1 Å).

Además, se evaluaron las conformaciones de las cadenas laterales calculando los ángulos de cadena lateral $\chi 1$ y $\chi 2$, y con estos valores se construyeron mapas de Janin, los cuales dan información sobre el movimiento/rotación de estas cadenas laterales. Estos valores se compararon con los valores experimentales obtenidos directamente de las estructuras del Protein Data Bank.

La Figura 32 muestra la colección de mapas de Janin para cada residuo del sitio activo, con excepción de la K201, la cual, como se mostró en la figura 21B está en una horquilla desordenada. Debido a su posición este residuo se mantiene la mayor parte del tiempo expuesta al solvente y, por tanto, muestrea todas las combinaciones permitidas para su cadena lateral en presencia o ausencia de ADN.

Estos mapas muestran que la R173 en el contexto de Cre libre (Figura 32A panel izquierdo) tiene una distribución centrada principalmente en [300°, 180°] (la conformación inicial), además de que puede muestrear fácilmente dos cuencas localizadas aproximadamente entre ~300° y 60° para $\chi 2$ mientras mantiene fijo el valor de $\chi 1$ en 300°. Cre^{Yin} y Cre^{Yang} (panel de en medio y panel derecho, respectivamente) tienen una distribución similar de valores. Sin embargo, en estos dos sistemas la cuenca localizada en [60, 180] para $\chi 1$ y $\chi 2$ está más poblada en comparación con Cre libre; este movimiento en $\chi 1$ podría alejar al NZ del fosfato escindible en *loxI*. E176 (Figura 32B) es un residuo de la segunda esfera, el cual no está en contacto directo con el ADN, pero se encarga de estabilizar a la R173[12]. Para este residuo en Cre libre la distribución de valores se mantiene cerca de la

inicial, teniendo cambios principalmente en $\chi 2$; este comportamiento de mantiene en los sistemas unidos a *loxI* (Cre^{Yin} y Cre^{Yang}).

H289 (Figura 32C) en Cre libre parece no mostrar ninguna preferencia por alguna de las combinaciones. Este residuo está localizado en el N-terminal de la hélice K, una región que es propensa a desplegarse, lo que podría explicar su variabilidad estructural. Sin embargo, Cre^{Yin} y Cre^{Yang} muestran una cuenca preferencialmente poblada en valores [300, 60] para $\chi 1$ y $\chi 2$, respectivamente, localizadas en el valor experimental para tres de las estructuras iniciales. Por otro lado, R292 (Figura 32D) muestrea valores localizados principalmente en [300°, 180°], como se observa en las estructuras experimentales en los tres sistemas comparados. El comportamiento de estos últimos dos residuos podría explicar porque es más fácil ensamblar un sitio activo funcional cuando *loxI* está presente (Figura 21A), ya que sus cadenas laterales no se desvían de las conformaciones iniciales.

W315 (Figura 32E) está restringido en valores cercanos a 300° para $\chi 1$ y $\chi 2$. Esto se debe principalmente a que es un residuo que realiza múltiples interacciones con los residuos vecinos que lo estabilizan en una conformación particular, cosa que sucede tanto en ausencia como en presencia de ADN.

Finalmente, el residuo catalítico Y324 (Figura 32F), muestrea dos conformaciones principales en $\chi 1$, localizadas en 300° (valor experimental) y en 180°, además de múltiples valores para $\chi 2$. Esto último sugiere que el anillo de este residuo es libre de rotar. La rotación de este anillo no altera la posición del grupo OH, sin embargo, cambios en el valor de $\chi 1$ sí afectan la posición del resto de la cadena lateral. Esto sucede principalmente en los sistemas Cre libre y Cre^{Yang}. En el caso de Cre^{Yin} (panel de en medio), estos valores están aún más restringidos a cambios en $\chi 2$ solamente. Este residuo, junto con H289, R292, y W315 podrían ser los principales participantes en la organización del sitio activo (Figura 33).

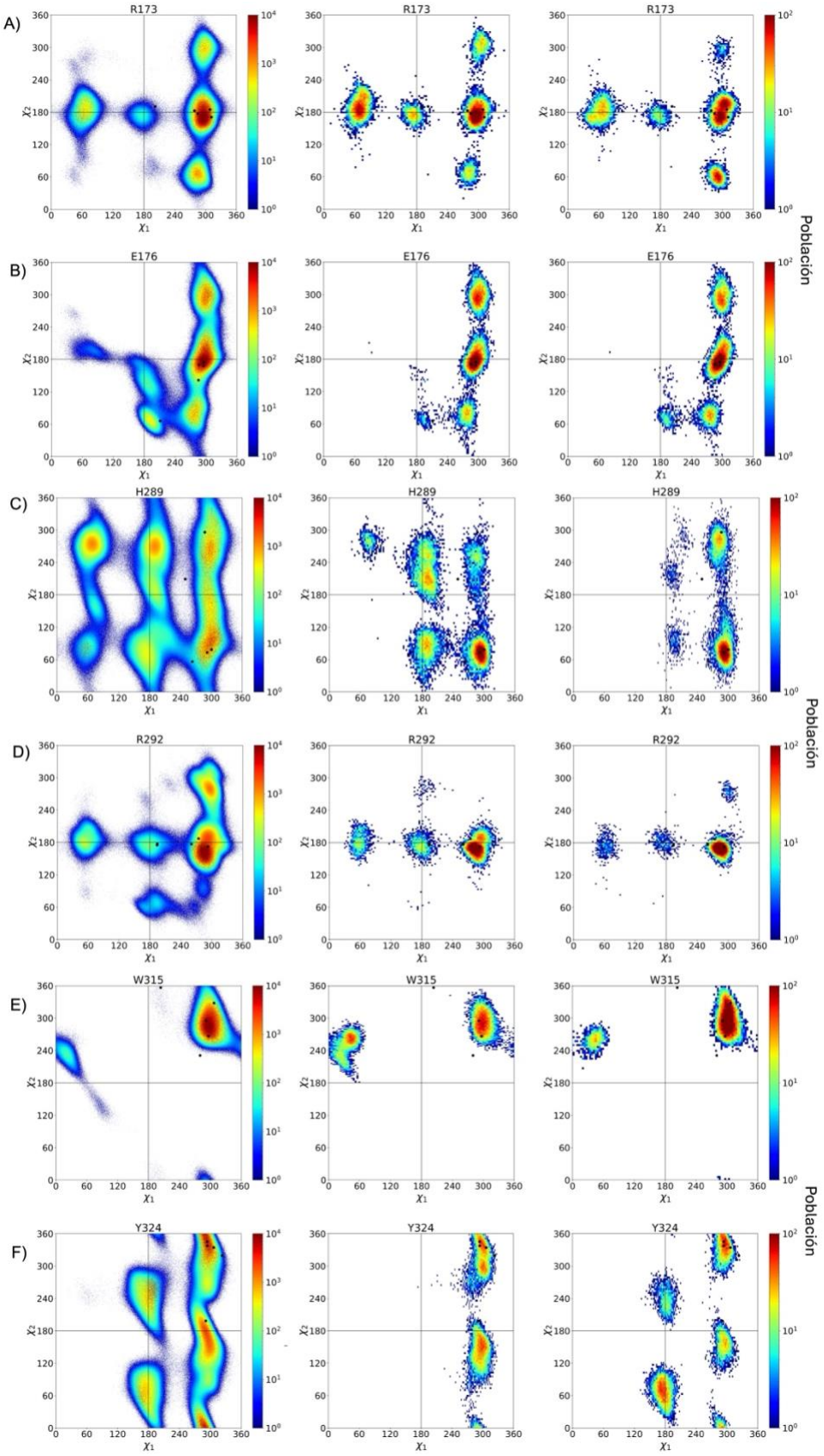

A)
R173
B)
E176
C)
H289
D)
R292
E)
W315
F)
Y324
χ_1
χ_2
Población

Figura 32. Distribución de ángulos diedros de cadena lateral. Mapas de Janin para cada residuo del sitio activo. Los valores iniciales se marcan con puntos negros dentro de cada mapa. La barra de color indica el número de miembros en cada intervalo de combinaciones; panel izquierdo Cre libre, panel central Cre^{Yin}, panel derecho Cre^{Yang}.

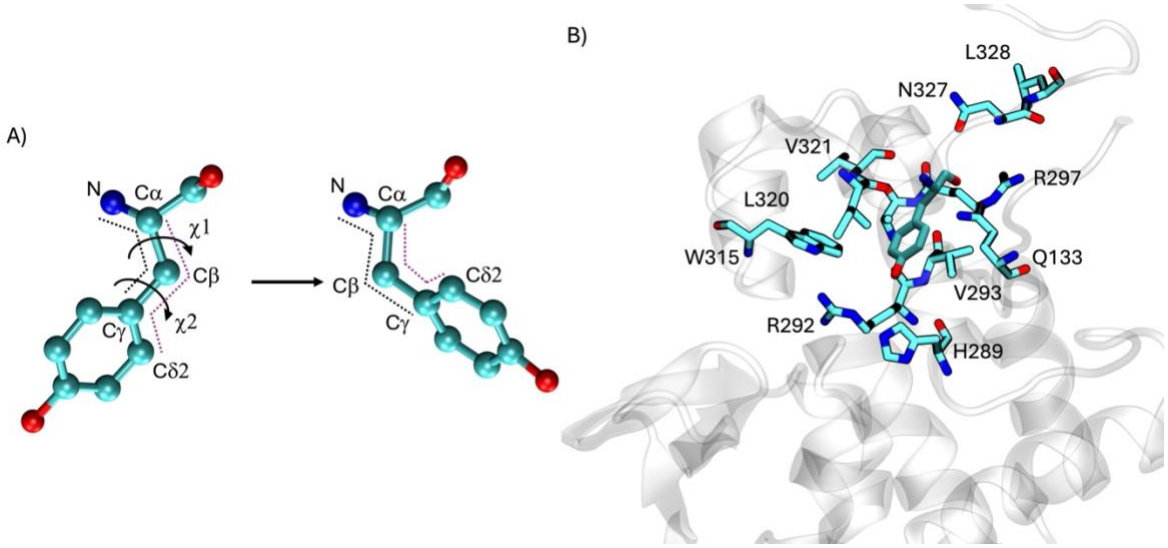


Figura 33. Definición de ángulos $\chi 1$ y $\chi 2$. A la derecha se muestran los átomos que definen al ángulo diedro $\chi 1$ (N-Cα-Cβ-Cγ, líneas puntadas negras) y a $\chi 2$ (Cα-Cβ-Cγ-C$\delta 2$, líneas punteadas moradas). Las flechas indican el enlace sobre el que se hace el giro/rotación. A la izquierda se muestra el cambio en estos ángulos llevando a una posición diferente del anillo de la tirosina. B) Detalle de la interacción de la Y324 con H289, R292 y W15, además se marcan los residuos cercanos a esta tirosina.

7.2.9 La hélice A muestrea múltiples estados

En el contexto de un intasoma, Cre realiza múltiples interacciones proteína-proteína que contribuyen a la estabilización del complejo. En ausencia de estas interacciones, Cre muestra un incremento en su flexibilidad.

Por otro lado, las estructuras disponibles para un monómero unido a un medio sitio *loxA* (PDB ID: 7RHY) y para un dímero unido a un sitio *loxI* completo (PDB ID: 7RHZ) muestran una posición diferente de la hélice A en comparación con la observada en un complejo tetramérico. En este último, dicha hélice interactúa con el dominio CBD del monómero adyacente mediante puentes salinos formados entre la hélice A y la hélice E.

En los complejos mencionados anteriormente, la hélice A está inclinada hacia el ADN, adoptando una conformación incapaz de interactuar con el monómero adyacente. Esto sugiere que el complejo requiere de interacciones cuaternarias para ensamblar un intasoma funcional.

Se evaluó la distancia y el ángulo formado entre las hélices A y B, con el fin de tener información sobre un posible mecanismo de ajuste inducido en los pasos finales del armado del complejo tetramérico. Este comportamiento se comparó con el sistema CBD-aislado y los monómeros Cre^{Yin} y Cre^{Yang} unidos a *loxI.*

La distribución de los sistemas Cre libre y CBD-aislado (Figura 34A y 34B) muestran dos cuencas bien pobladas, una en valores positivos correspondiente a las estructuras tipo monómero (140°, 10.1 Å) y dímero (141°, 10 Å y 130°, 10.6 Å, para Cre^{Yin} y Cre^{Yang}, respectivamente) y otra en valores negativos que corresponde a los valores de un monómero activo (-149°, 13.2 Å) y otro inactivo (154°, 13.5 Å) en el contexto de un intasoma. Estos dos estados están conectados por conformaciones intermedias; estas conformaciones muestran que la hélice A puede ir de un estado otro en un movimiento tipo limpiaparabrisas. Por otro lado, las simulaciones del sistema Cre_2-*loxI* (Figura 33¿4C y 34D) se quedan cerca de los valores iniciales, con excepción de Cre^{Yang} que en algunos casos logra visitar estados tipo intasoma. Esto puede ser debido a la presencia de R24, K25 y R32 en la hélice A que podrían promover preferencialmente su interacción con el ADN, así como interacciones con la hélice B que la estabilizan en esta posición.

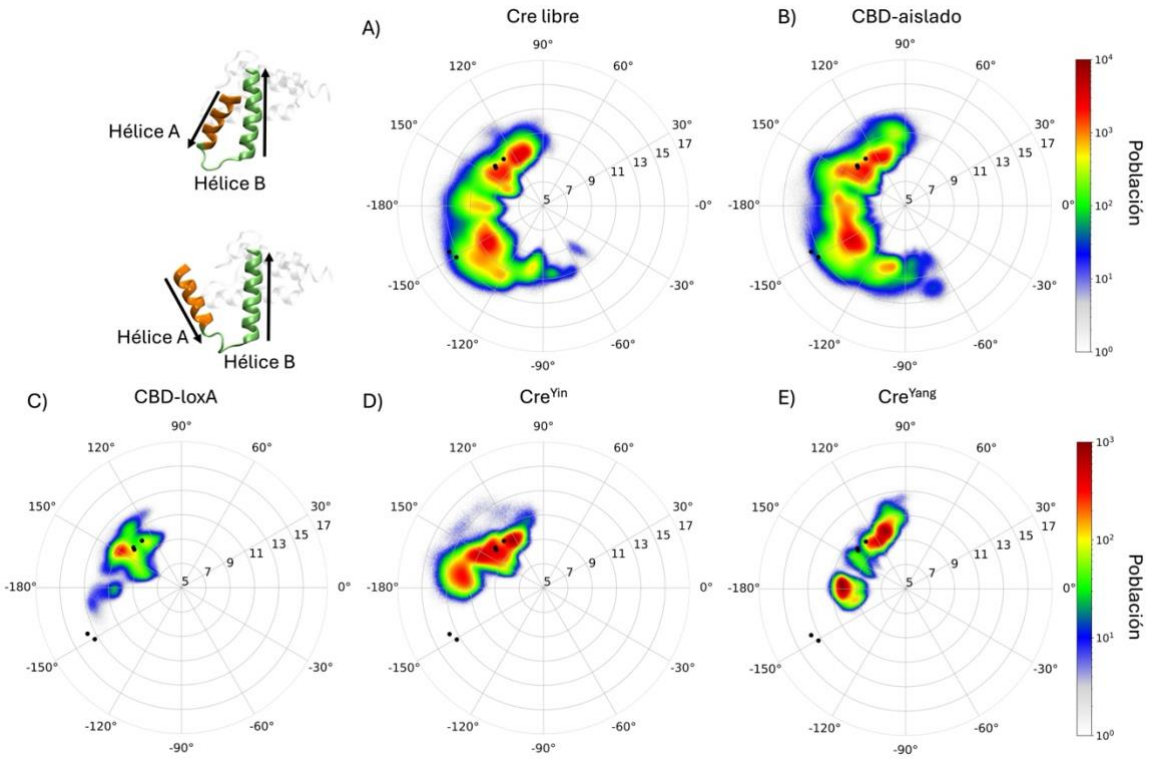


Figura 34. Paisaje conformacional de la hélice A en diferentes contextos. Histogramas en 2D para cada combinación de ángulos y distancias entre las hélices A y B para A) Cre libre, B) CBD aislado, C) CBD-*loxA*, D) Cre^{Yin} y E) Cre^{Yang}. D y E se muestran también en la Figura 10. La escala de color muestra el número de miembros en cada bin. En panel A muestra las dos posiciones de la hélice A en el contexto de un monómero/dímero y de un intasoma. Los valores experimentales se muestran como puntos negros en cada paisaje.

Sin embargo, a pesar de que la hélice A puede visitar los ángulos correspondientes a una conformación tipo intasoma, la distancia entre las hélices A y B sigue siendo más corta comparada con las estructuras experimentales (puntos negros en cada paisaje). Este comportamiento sucede en todos los sistemas analizados, sugiriendo que se necesita de un ajuste extra una vez que el complejo se formó para mantener a la hélice A en la posición adecuada (Figura 35).

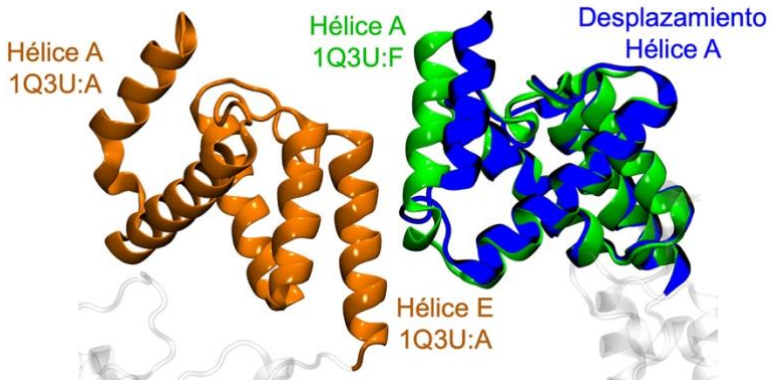


Figura 35. Desplazamiento de la hélice A. Se muestra en diferente color una cadena activa (anaranjado) y una cadena inactiva (verde) obtenidas de la estructura cristalográfica 1Q3U. En azul se muestra el sobrelape de la estructura con el valor más poblado de ángulo y distancia muestreado a lo largo del tiempo de simulación para Cre libre.

Debido a que los sistemas unidos a ADN muestran un movimiento restringido por la interacción de la hélice A con el ADN, contamos el número de transiciones de los sistemas Cre libre y CBD aislado, entre los dos estados principales (monómero/dímero e intasoma). Para esto definimos un nuevo estado como el estado “Barrera” en el cual la hélice A puede pasar de un estado tipo monómero/dímero o intasoma hacía este nuevo estado y, a partir de este, saltar hacia el estado vecino o regresar al estado original. También se definió estado “desordenado” en el que están todas las estructuras en las que la hélice A está desplegada y/o doblada. La Figura 36 muestra estas transiciones.

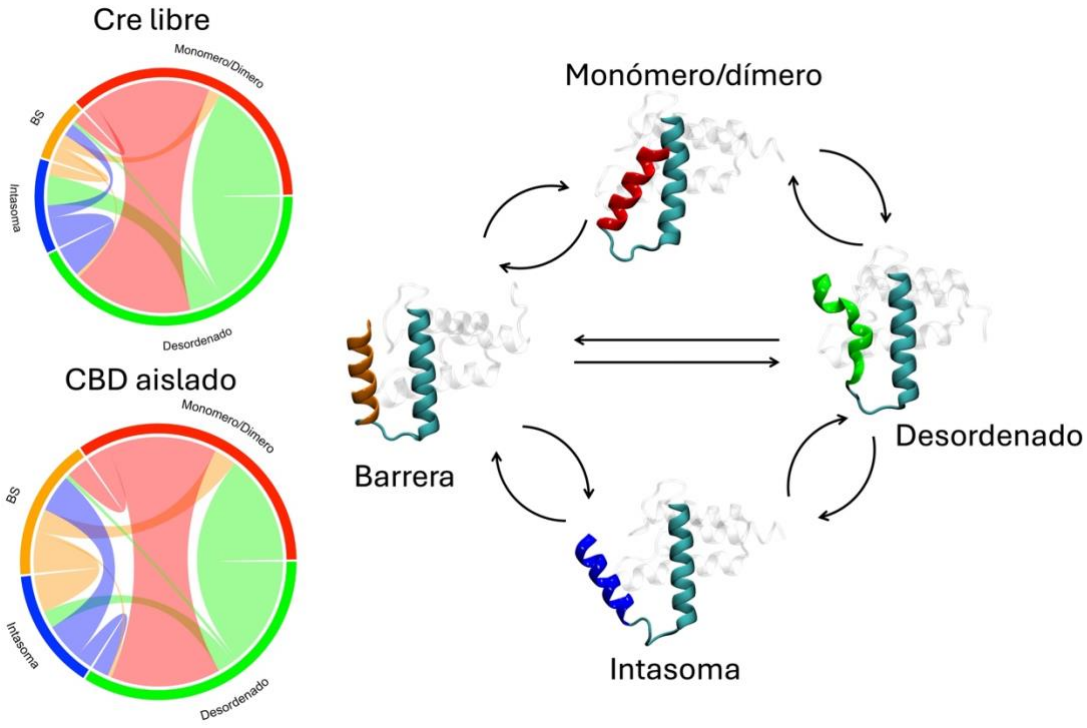


Figura 36. Transiciones entre estados. A la izquierda se muestra el gráfico de transiciones mostrando los saltos de un estado a otro y la conectividad entre ellos. A la derecha se muestran estructuras representativas de cada uno de los estados muestreados por la hélice A respecto a la hélice B (cian). Las flechas entre cada estructura muestran el flujo entre los saltos.

En ambos sistemas sin ADN, la hélice A muestrea todos los estados sin importar la estructura inicial. Estos estados están conectados entre sí a través del estado barrera o por el estado desordenado, siendo este último el único estado que se conecta a todos los demás. El movimiento tipo “limpia parabrisas” mostrado por la hélice A, así como el hecho de que se alcanzan estructuras tipo intasoma, sugiere un mecanismo de selección de conformación durante el armado del tetrámero. Además, este movimiento es independiente de la presencia del dominio CAT, ya que en el sistema CBD aislado las transiciones son similares a las de Cre libre. En los sistemas unidos a *loxI* la hélice A también tiene eventos de desplegamiento y plegamiento, sugiriendo un comportamiento flexible.

7.2.10 “Bisagra” del dominio CAT

El dominio CAT tiene una cavidad formada por las hélices I y K. En el contexto de un dímero, sólo una de estas cavidades está ocupada por la hélice N del monómero adyacente (Figura 11), dado que el sitio de unión en el otro monómero fue mutado para evitar la interacción. Los resultados mostrados para el dímero silvestre mostraron que la hélice N puede formar un complejo en el cual las dos cavidades están ocupadas. En el contexto del intasoma, la hélice N de un monómero interactúa con la cavidad de la subunidad adyacente en el mismo sitio *loxP*, del mismo modo, esta subunidad interactúa con la cavidad del monómero adyacente en un sitio *loxP* diferente (Figura 5, paso 4).

Se evaluó la apertura y cierre de esta cavidad en el resto de los sistemas que tienen al dominio CAT (Cre libre, CAT-delN y CAT). Estos resultados se compararon con el comportamiento antes visto con los sistemas unidos al ADN. Para poner hacer una comparación directa, todos los paisajes conformacionales se muestran en la misma escala (Figura 37).

Los resultados muestran que los sistemas unidos a ADN no muestrean valores amplios de ángulos y distancias, manteniéndose cerca de los valores iniciales. Este comportamiento está dado por interacciones proteína-proteína entre monómeros y proteína-ADN que le restan flexibilidad a Cre. Por otro lado, el sistema Cre libre muestra la distribución más amplia de ángulos comparado con los otros dos sistemas. Interesantemente, el sistema CAT-delN muestra una disminución en la distancia, pero muestrea valores de ángulos hacia 90°. Finalmente, el sistema CAT muestra un paisaje conformacional restringido, manteniéndose cerca de los valores iniciales. Como se mostró en secciones previas, tanto el conector interdominios como el extremo C-terminal interactúan con la cavidad evaluada y con los residuos del sitio activo, los cuales se localizan detrás de este subdominio. Estas interacciones podrían explicar la disminución de movimiento, y sugieren que estas interacciones suceden durante la mayor parte del tiempo de simulación.

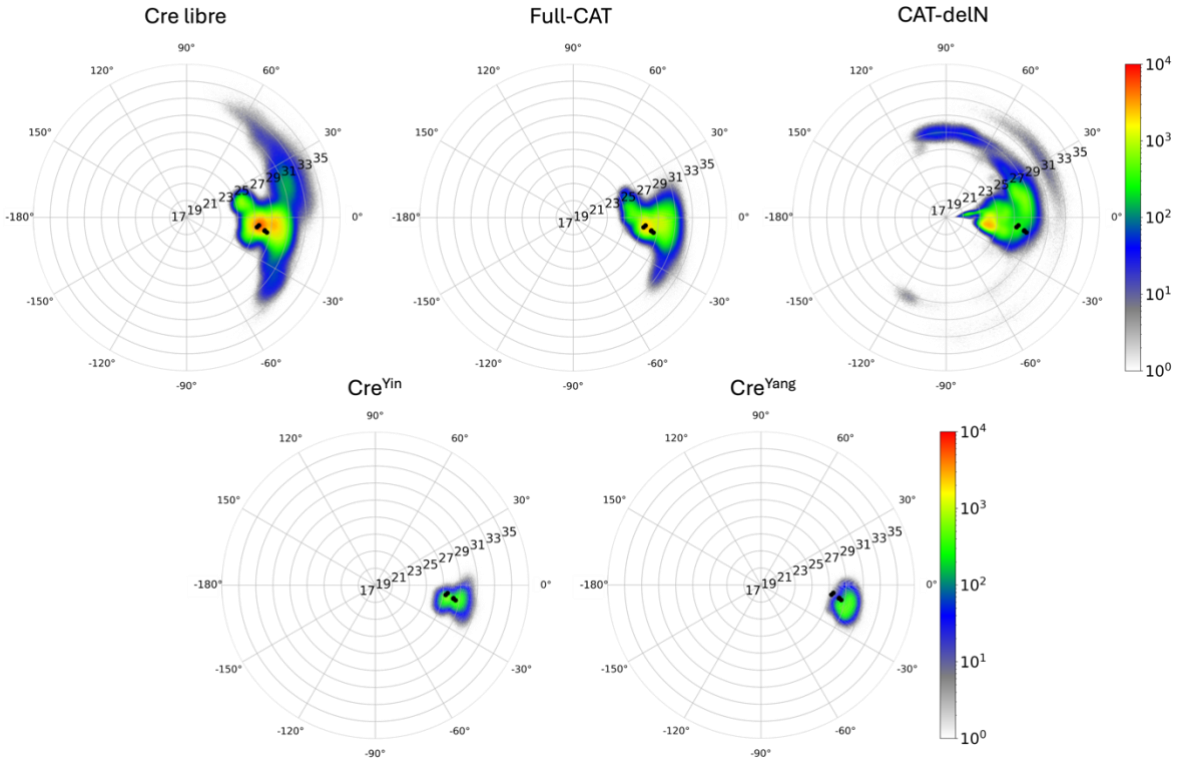


Figura 37. Paisajes conformaciones de las hélices I-K para todos los sistemas con dominio CAT. Los valores iniciales se muestran como puntos negros en cada paisaje.

7.2.11 Estructura de CBD aislado y CBD-*loxA*

Entre los estudios experimentales hechos para el complejo CBD unido a una horquilla de ADN con la secuencia *loxA* (CBD-*loxA*) se encuentra el realizado por la Dra. Leidys French Pacheco en su tesis doctoral[66]. Sus resultados mostraron que el dominio CBD tiene una IDR en el extremo N-terminal, la cual no sufre un cambio de desplegado a plegado en el complejo con *loxA*.

Calculamos el espectro de dicroísmo circular (CD, por sus siglas en inglés) de los sistemas CBD aislado y CBD-*loxA* con el software SESCA[67], y lo comparamos con el espectro experimental obtenido en la tesis antes mencionada. También se obtuvo el promedio de estructura tipo alfa hélice para los dos sistemas. La Figura 38 muestra que el patrón de CD se conserva y es comparable con el obtenido experimentalmente: se muestran los dos mínimos en 208 y 220 nm, mínimos típicos

de conformaciones helicoidales. Por otro lado, los porcentajes de estructura tipo alfa hélice para el sistema CBD aislado y CBD-loxA difieren en 1.7% entre sí. Como se mencionó anteriormente, se encontró que la unión a ADN promueve un cambio conformacional en ambos dominios, pero el porcentaje de estructura tipo alfa hélice no variaba. Es importante notar que el porcentaje de hélice alfa obtenido en las simulaciones es menor comparado con el obtenido experimentalmente (38%). Sin embargo, los valores se mantienen dentro de la misma escala, sugiriendo que este dominio libre o unido a ADN, se mantiene plegado y contiene múltiples elementos helicoidales.

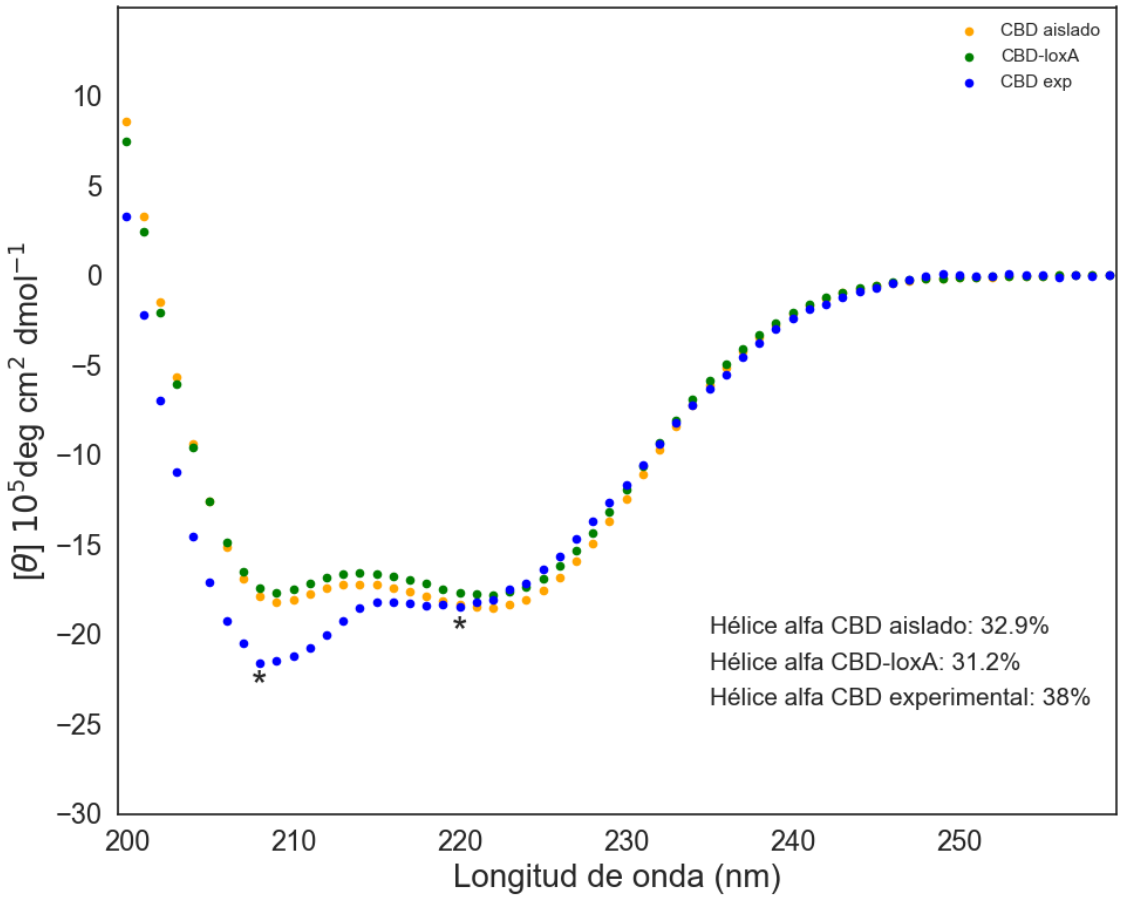


Figura 38. Espectro de dicroísmo circular. Comparación de los espectros de dicroísmo circular experimental (azul), CBD aislado (anaranjado) y CBD-*loxA*

(verde). Los asteriscos muestran los dos mínimos asociados a la formación de estructuras helicoidales (208 y 220 nm).

Además de determinar el porcentaje de estructura helicoidal, también se concluyó que el dominio CBD tiene una IDR en el extremo N-terminal de aproximadamente 40 residuos[66]. Estos residuos que se predicen como desordenados incluyen a la hélice A, segmento que previamente se mostró tiene un movimiento tipo limpia parabrisas que le permite visitar múltiples estados, entre ellos los correspondientes a un monómero/dímero e intasoma.

Para evaluar la formación y ruptura de estructura secundaria de este segmento, se calculó el porcentaje hélice (suma de las contribuciones de estructura tipo hélice alfa, hélice 3_{10} y hélice pi) a lo largo del tiempo para todos los sistemas (Anexo IV). La Figura 39 muestra el cambio de estructura secundaria para el sistema proveniente del monómero con y sin ADN para cada réplica a lo largo del tiempo de simulación. En ambos sistemas, la hélice A pierde y gana estructura helicoidal, y esto ocurre independientemente de que haya o no ADN. Sin embargo, la presencia de *loxA* sí tiene un efecto sobre el movimiento antes descrito para la hélice A. Como se discutió en secciones previas, esta restricción de movimiento puedes explicarse con interacciones electrostáticas realizadas entre residuos positivos y el ADN o mediados por agua.

Por otro lado, el comportamiento dinámico de la hélice A con y sin ADN correlaciona con los resultados que muestran que la unión de CBD a *loxA* no tiene un efecto sobre el plegamiento de esta hélice, pero sí forma un complejo estable[66].

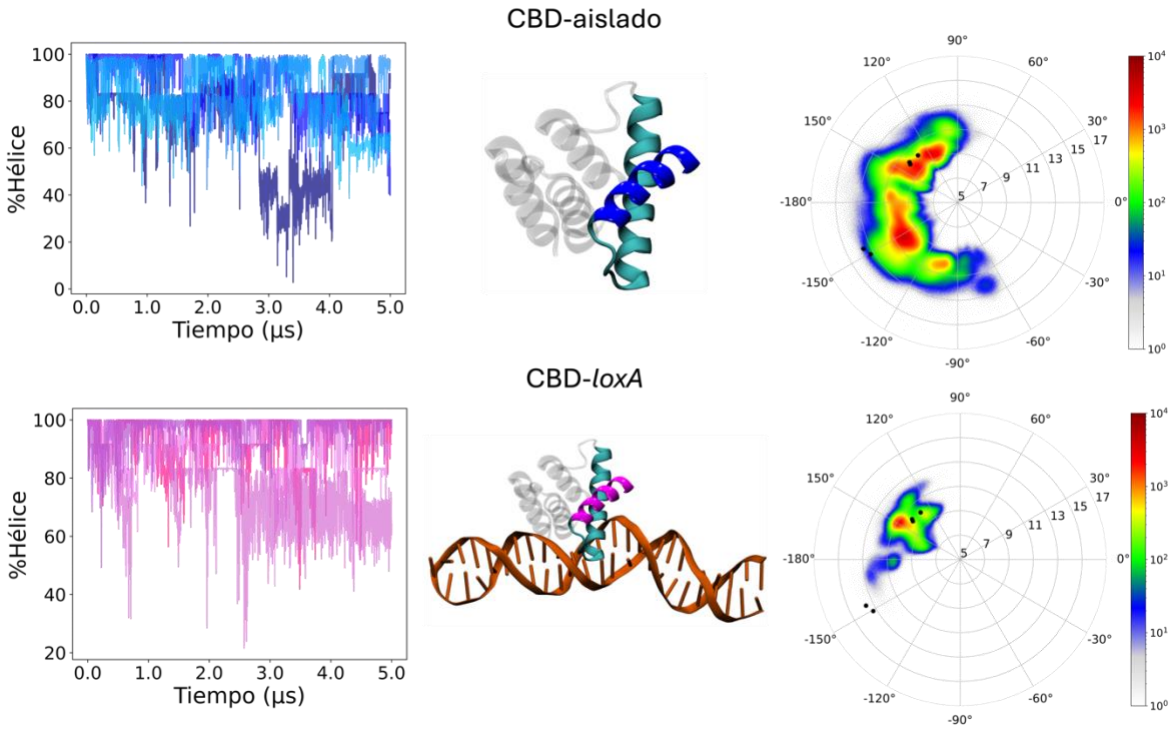


Figura 39. Estructura secundaria de la hélice A con y sin ADN. El panel izquierdo muestra el trazo de estructura secundaria (hélice) de la hélice A para cada réplica a lo largo del tiempo de simulación. El panel central muestra la estructura del sistema usado para cada cálculo. El panel derecho muestra la distribución de ángulos y distancias de la hélice A respecto a la hélice B (tomado de la Figura 23) para comparación.

8. Conclusiones

En el contexto del dímero unido a *loxI*, la hélice N no tiene un efecto claro sobre el cambio del valor de *roll*. Sin embargo, durante el tiempo de simulación se logró la obtención de un intermediario similar al propuesto por Stachowski y colaboradores[14] en el cual hay un intercambio de hélices y, por tanto, un cambio en la dirección de la curvatura de *loxP,* este cambio ocurrió a nivel del valor de *roll* en la posición 7 del espaciador, llevando a una estructura parecida a la presente en el intasoma tipo 1Q3U.

En el contexto del dímero, la hélice A se mantiene principalmente cerca de la conformación inicial, los valores de ángulo y distancia muestran que incluso puede inclinarse más hacia el ADN. Esto se puede explicar con las cargas positivas de la hélice A que la acercan hacia *loxI*. Esta conformación oculta una cavidad hidrofóbica formada entre las hélices A, B y C, cavidad que se ha sugerido es estabilizada por interacciones con el extremo N-terminal del dominio CBD[41]; residuos cuya estructura no ha sido determinada experimentalmente. Por otro lado, no se observaron conformaciones tipo intasoma en ninguno de los dos monómeros en la simulación con *loxI*. Esto sugiere que se necesitan de interacciones cuaternarias favorables que le permitan a esta hélice alejarse del ADN e interactuar con el monómero adyacente, hablando de un posible mecanismo de ajuste inducido.

En ausencia de ADN, Cre es una proteína flexible. Esto se ve reflejado en los histogramas de distribución de radio de giro y la distancia entre extremos. Los dominios son libres de rotar uno respecto a otro (paisajes conformacionales construidos con las hélices D y G), este comportamiento podría ser importante en los primeros pasos de unión a *loxP*, en el cual la flexibilidad intrínseca de la proteína es clave para la selección y unión al sitio blanco; esta libre rotación está restringida en los sistemas unidos a *loxI*. A pesar de la flexibilidad, ambos dominios se mantienen plegados, lo que concuerda con estudios previos que muestran que los

dominios tienen plegamiento independiente[11] además de tener pocos contactos entre ellos.

El movimiento de la hélice A parece ser uno de los últimos pasos durante la formación del intasoma. Esta hélice, en ausencia de ADN, es capaz de muestrear conformaciones tipo monómero/dímero e intasoma, sin importar el punto de partida. Este comportamiento se mantiene en los sistemas Cre libre y CBD-aislado. Además de exhibir este movimiento, esta hélice también tiene eventos de desplegamiento y plegamiento. El comportamiento dinámico de este segmento pudiera explicar por qué los primeros 40 residuos de este dominio han sido predichos como una región intrínsecamente desordenada[66], además de que los primeros 19 no han sido resueltos en las estructuras experimentales disponibles.

El dominio CAT tiene una “bisagra” formada por las hélices I-K. Esta cavidad tiene un movimiento tipo almeja que abre y cierra la región de interacción con la hélice N, facilitando o impidiendo la interacción entre estos dos segmentos. En los sistemas unidos a ADN, el movimiento de esta cavidad está impedido por interacciones con el ADN y con el dominio CAT adyacente.

La hélice N es un segmento desordenado que visita múltiples conformaciones, lo que le permite realizar múltiples contactos con el resto de Cre. La interacción de este segmento con los residuos del sitio activo lleva a la formación de un complejo auto-inhibitorio. En Cre libre esta conformación es poco persistente, probablemente debido a interacciones con el dominio CBD que podría estar interponiéndose entre la hélice N y el sitio activo. Por otro lado, en el sistema CAT, este complejo tiene una mayor persistencia, confirmando que la presencia del CBD sí podría estar interfiriendo con esta interacción. El conector flexible también interactúa con el sitio activo y, en menor medida, con el resto de Cre. La conformación auto-inhibitoria fue propuesta previamente para la interacción de la hélice N con el sitio activo[11]. Sin embargo, no se ha establecido un papel auto-inhibitorio similar para el conector interdominios. Las interacciones realizadas entre el conector y la hélice N con el sitio activo podría indicar que existe más de una forma de inhibir a Cre en ausencia

de ADN. Además, a diferencia del sistema dímero unido a *loxI* y CAT, la cavidad I-K del sistema CAT-delN muestrea más conformaciones, sugiriendo que la ausencia del conector y la hélice N lleva al colapso de la cavidad (muestrea distancias más cortas) pero a una amplia distribución de ángulos. En el sistema CAT, las interacciones del conector y de la hélice N podrían estar promoviendo la disminución del movimiento de esta cavidad.

La flexibilidad observada en los sistemas Cre libre, CBD aislado, CAT-delN y CAT muestran que, en presencia de ADN, Cre se encuentra en una conformación restringida, la cual difiere a su estructura cuando está libre. Esta conformación restringida está estabilizada por múltiples interacciones proteína-proteína y proteína-ADN, las cuales, van apareciendo de forma gradual a lo largo del ensamblado del intasoma. Además, el aumento en flexibilidad de Cre libre visto en las simulaciones explica el porqué no se ha podido obtener una estructura experimental de este sistema.

El tiempo de simulación empleado es suficiente para capturar la dinámica de Cre en solución y de sus dominios por separado, ya que la comparación con datos experimentales dio una buena correlación.

9. Perspectivas

- Hacer simulaciones del dímero unido a *loxP* haciendo mutaciones en las bases correspondientes al espaciador, esto con el fin de establecer una relación entre la secuencia de esta región y la facilidad de doblar al ADN hacía una u otra dirección.

- Completar el set de simulaciones del sistema CAT-delN.

- Se cuenta con las simulaciones del dominio CBD unido a *loxA*; hacer las simulaciones de Cre completa y el dominio CAT unidos a diferentes secuencias *lox,* podría dar información sobre las interacciones iniciales en el proceso de colonización de la secuencia *lox*, así como de la estabilidad de los complejos.

- Determinar nuevas coordenadas de reacción que nos permitan caracterizar de forma más detallada los cambios conformacionales de Cre asociados a la unión a *lox* así como determinar si esta secuencia tiene un efecto sobre dichas conformaciones.

- Optimizar las herramientas de análisis existentes que permitan hacen una comparación directa con datos experimentales.

- Dado que se ha observado que el extremo N-terminal de Cre es un segmento desordenado, hacer simulaciones del dominio CBD completo usando estrategias de muestreo acelerado podrían dar indicios de la posición de esta región en diferentes pasos del ensamblado del intasoma.

- Proponer experimentos de FRET que nos permitan corroborar los resultados obtenidos con las distancias entre las hélices D y G. Por otro lado, experimentos de Dispersión de Rayos X a Bajo Ángulo (SAXS) o Dispersión Dinámica de Luz (DLS), podrían ser de utilidad para validar los resultados de radio de giro.

Este trabajo se ha presentado en sus diferentes versiones en múltiples congresos nacionales e internacionales. A continuación, se enlistan algunos de ellos:

1. Effect of the N helix on the structure and dynamics of the catalytic domain of Cre recombinase. Congreso Internacional de Biología Estructural. Estructura, Dinámica e Interacciones. 28 de septiembre al 2 de octubre 2020.
2. Dynamical behavior of the CAT domain from Cre recombinase, in the absence of and with the *cis* interaction of the N helix. 65th Annual Meeting of the Biophysical Society. 22 - 26 de febrero 2021, evento virtual.
3. Essential dynamics of the catalytic core of the Cre family of recombinases: comparison of Cre and FLP homologs. LatinXchem, 20 septiembre 2021. Twitter (ahora X).
4. Characterizing the energy landscape of Cre recombinase catalytic domain through molecular dynamics simulations. 16 al 20 de mayo 2022, Hamburgo, Alemania.
5. Interdomain crosstalk in Cre recombinase. BioPhysMex 2022. 10 de noviembre de 2022, Ciudad de México.
6. Multiple conformations of Helix A from the N-terminal domain of Cre connect those in monomer/dimers and in intasomes. 37th Annual meeting of Protein Society. 13 al 16 de julio 2023, Boston, Massachusetts.
7. Multiple conformations of helix A from Cre recombinase are context-dependent. LatinXchem, 16 y 17 de octubre 2023, Twitter (ahora X).
8. Hélice A de la recombinasa Cre como interruptor molecular en el armado del intasoma. Current biophysics topics in Mexico, 4 abril 2024, Instituto de Ciencias Físicas, UNAM, Cuernavaca, Morelos, México.

9. Helix A from Cre recombinase as a molecular switch in intasome assembly. 23 al 26 de julio 2024, Vancouver, Canadá.
10. Flexibility abd self-inhibition: exploring the conformational diversity of Cre recombinase. LatinXchem, 14 al 15 de octumbre 2024, Twitter (ahora X).

También se han presentado talleres de divulgación con el título “Corte y Confección de Cromosomas”, eventos en los que se explica la reacción de recombinación hecha por Cre. Este taller se ha presentado en:

1. Rally Científico 2023, Museo de Ciencias de Morelos.
2. Museo Fest 2025, Museo de Ciencias de Morelos.

Además, se escribió una nota de divulgación para el Diario de Morelos, titulada “Personalidad dinámica de la recombinasa Cre, un arquitecto molecular”. Esta publicación puede ser encontrada entrando al siguiente enlace:

https://acmor.org/publicaciones/la-personalidad-din-mica-de-la-recombinasa-cre-un-arquitecto-molecular

10. Bibliografía

(1) Nagy, A. Cre Recombinase: The Universal Reagent for Genome Tailoring. *genesis* **2000**, *26* (2), 99–109. https://doi.org/10.1002/(SICI)1526-968X(200002)26:2<99::AID-GENE1>3.0.CO;2-B.

(2) Meinke, G.; Bohm, A.; Hauber, J.; Pisabarro, M. T.; Buchholz, F. Cre Recombinase and Other Tyrosine Recombinases. *Chem. Rev.* **2016**, *116* (20), 12785–12820. https://doi.org/10.1021/acs.chemrev.6b00077.

(3) Van Duyne, G. D. Cre Recombinase. *Microbiol. Spectr.* **2015**, *3* (1). https://doi.org/10.1073/pnas.1304749110.

(4) Grindley, N. D. F.; Whiteson, K. L.; Rice, P. A. Mechanisms of Site-Specific Recombination. *Annu. Rev. Biochem.* **2006**, *75* (1), 567–605. https://doi.org/10.1146/annurev.biochem.73.011303.073908.

(5) Abremski, K.; Hoess, R. Bacteriophage P1 Site-Specific Recombination. *J. Biol. Chem.* **1983**, *259* (3), 1509–1514. https://doi.org/10.1016/S0021-9258(17)43437-5.

(6) Foster, M. P.; Benedek, M. J.; Billings, T. D.; Montgomery, J. S. Dynamics in Cre-loxP Site-Specific Recombination. *Curr. Opin. Struct. Biol.* **2024**, *88*, 102878. https://doi.org/10.1016/j.sbi.2024.102878.

(7) Łobocka, M. B.; Rose, D. J.; Plunket, G.; Rusin, M.; Samojedny, A.; Lehnherr, H.; Yarmolinsky, M. B.; Blattner, F. R. Genome of Bacteriophage P1. *J. Bacteriol.* **2004**, *186* (21), 7032–7068. https://doi.org/10.1128/JB.186.21.7032–7068.2004.

(8) Hoess, R. H.; Ziese, M.; Sternberg, N. P1 Site-Specific Recombination: Nucleotide Sequence of the Recombining Sites. *Proc. Natl. Acad. Sci.* **1982**, *79* (11), 3398–3402. https://doi.org/10.1073/pnas.79.11.3398.

(9) Sternberg, N.; Hamilton, D. Bacteriophage P1 Site-Specific Recombination I. Recombination Between loxP Sites. *J. Mol. Biol.* **1981**, *150* (4), 467–486. https://doi.org/10.1016/0022-2836(81)90375-2.

(10) Sternberg, N.; Hamilton, D.; Austin, S.; Yarmolinsky, M.; Hoess, R. Site-Specific Recombination and Its Role in the Life of Bacteriophage P1. *Cold Spring Harb. Symp. Quant. Biol.* *45* (1), 297–309. https://doi.org/10.1101/sqb.1981.045.01.042.

(11) Unnikrishnan, A.; Amero, C.; Yadav, D. K.; Stachowski, K.; Potter, D.; Foster, M. P. DNA Binding Induces a *Cis* -to- *Trans* Switch in Cre Recombinase to Enable Intasome Assembly. *Proc. Natl. Acad. Sci.* **2020**, *117* (40), 24849–24858. https://doi.org/10.1073/pnas.2011448117.

(12) Gibb, B.; Gupta, K.; Ghosh, K.; Sharp, R.; Chen, J.; Van Duyne, G. D. Requirements for Catalysis in the Cre Recombinase Active Site. *Nucleic Acids Res.* **2010**, *38* (17), 5817–5832. https://doi.org/10.1093/nar/gkq384.

(13) Ennifar, E. Crystal Structure of a Wild-Type Cre Recombinase-loxP Synapse Reveals a Novel Spacer Conformation Suggesting an Alternative Mechanism for DNA Cleavage Activation. *Nucleic Acids Res.* **2003**, *31* (18), 5449–5460. https://doi.org/10.1093/nar/gkg732.

(14) Stachowski, K.; Norris, A. S.; Potter, D.; Wysocki, V. H.; Foster, M. P. Mechanisms of Cre Recombinase Synaptic Complex Assembly and Activation Illuminated by Cryo-EM. *Nucleic Acids Res.* **2022**, *50* (3), 1753–1769. https://doi.org/10.1093/nar/gkac032.
(15) Mack, A.; Sauer, B.; Abremski, K.; Hoess, R. Stoichiometry of the Cre Recombinase Bound to the *Lox* Recombining Site. *Nucleic Acids Res.* **1992**, *20* (17), 4451–4455. https://doi.org/10.1093/nar/20.17.4451.
(16) Rufer, A. W. Non-Contact Positions Impose Site Selectivity on Cre Recombinase. *Nucleic Acids Res.* **2002**, *30* (13), 2764–2771. https://doi.org/10.1093/nar/gkf399.
(17) Eroshenko, N.; Church, G. M. Mutants of Cre Recombinase with Improved Accuracy. *Nat. Commun.* **2013**, *4* (1), 2509. https://doi.org/10.1038/ncomms3509.
(18) Hartung, M.; Kisters-Woike, B. Cre Mutants with Altered DNA Binding Properties. *J. Biol. Chem.* **1998**, *273* (36), 22884–22891. https://doi.org/10.1074/jbc.273.36.22884.
(19) Baldwin, E. P.; Martin, S. S.; Abel, J.; Gelato, K.; Kim, H.; Schultz, P.; Santoro, S. A Specificity Switch in Selected Cre Recombinase Variants Is Mediated by Macromolecular Plasticity and Water. *J. Chem. Biol.* **2003**, *10*, 1085–1094. https://doi.org/10.1016/j .c he m bi ol . 20 03 . 10 .0 1 5.
(20) Van Duyne, G. D. A Structural View of Cre- *loxP* Site-Specific Recombination. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30* (1), 87–104. https://doi.org/10.1146/annurev.biophys.30.1.87.
(21) Wu, J.; Wang, M.; Yang, X.; Yi, C.; Jiang, J.; Yu, Y.; Ye, H. A Non-Invasive Far-Red Light-Induced Split-Cre Recombinase System for Controllable Genome Engineering in Mice. *Nat. Commun.* **2020**, *11* (1), 3708. https://doi.org/10.1038/s41467-020-17530-9.
(22) Rongrong, L.; Lixia, W.; Zhongping, L. Effect of Deletion Mutation on the Recombination Activity of Cre Recombinase. *Acta Biochim. Pol.* **2005**, *52* (2), 541–544. https://doi.org/10.18388/abp.2005_3470.
(23) Hollingsworth, S. A.; Dror, R. O. Molecular Dynamics Simulation for All. *Neuron* **2018**, *99* (6), 1129–1143. https://doi.org/10.1016/j.neuron.2018.08.011.
(24) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30* (10), 1545–1614. https://doi.org/10.1002/jcc.21287.
(25) Salomon-Ferrer, R.; Case, D. A.; Walker, R. C. A Good General Overview of the Amber Biomolecular Simulation Package. *WIREs Comput Mol Sci* **2013**, *3*, 198–210.

(26) Christen, M.; Hunenberger, H.; Bakowies, R.; Baron, R.; Burgi, R.; Geerke, D.; Heinz, T.; Kastenholz, M.; Krautler, V.; Oostenbrink, C.; Peter, C.; Trzesniak, D.; Van Gusteren, F. The GROMOS Software for Biomolecular Simulation: GROMOS05. *J Comput Chem 26*, 1719–1751. https://doi.org/10.1002/jcc.20303.
(27) Robertson, M.; Tirado-Rives, J.; Jorgensen, W. Improved Peptide and Protein Torsional Energetics with the OPLS-AA Force Field. *J Chem Theory Comput* **2015**, *11*, 3499–3509. https://doi.org/10.1021/acs.jctc.5b00356.
(28) Shen, W.; Zhou, T.; Shi, X. Enhanced Sampling in Molecular Dynamics Simulations and Their Latest Applications -A Review. *Nano Res* **2023**, *16*, 13474–13497. https://doi.org/10.1007/s12274-023-6311-9.
(29) Wang, Y.; Lamim Ribeiro, J. M.; Tiwary, P. Machine Learning Approaches for Analyzing and Enhancing Molecular Dynamics Simulations. *Curr. Opin. Struct. Biol.* **2020**, *61*, 139–145. https://doi.org/10.1016/j.sbi.2019.12.016.
(30) Stevens, J. A.; Grünewald, F.; van Tilburg, P. A. M.; König, M.; Gilbert, B. R.; Brier, T. A.; Thornburg, Z. R.; Luthey-Schulten, Z.; Marrink, S. J. Molecular Dynamics Simulation of an Entire Cell. *Front. Chem.* **2023**, *11*. https://doi.org/10.3389/fchem.2023.1106495.
(31) Dommer, A. C.; Wauer, N. A.; Angle, K. J.; Davasam, A.; Rubio, P.; Luo, M.; Morris, C. K.; Prather, K. A.; Grassian, V. H.; Amaro, R. E. RevealingtheImpactsofChemicalComplexityonSubmicrometer SeaSprayAerosolMorphology. *ACS Cent. Sci.* **2023**, *9* (6), 1088−1103. https://doi.org/10.1021/acscentsci.3c0018.
(32) Guo, F.; Gopaul, D.; Van Duyne, G. D. Structure of Cre Recombinase Complexed with DNA in a Site- Specific Recombination Synapse. *Nature* **1997**, *389* (4), 40–46. https://doi.org/10.1038/37925.
(33) Ghosh, K.; Lau, C.; Guo, F.; Segall, A.; Van Duyne, G. D. Peptide Trapping of the Holliday Junction Intermediate in Cre-loxP Site-Specific Recombination. *J. Biol. Chem. 280* (9), 8290–8299. https://doi.org/10.1074/jbc.M411668200.
(34) Woods, K. C.; Martin, S. S.; Chu, V. C.; Baldwin, E. P. Quasi-Equivalence in Site-Specific Recombinase Structure and Function: Crystal Structure and Activity of Trimeric Cre Recombinase Bound to a Three-Way Lox DNA Junction. *J. Mol. Biol.* **2001**, *313* (1), 49–69. https://doi.org/10.1006/jmbi.2001.5012.
(35) Martin, S. S.; Pulido, E.; Chu, V. C.; Lechner, T.; Baldwin, E. P. The Order of Strand Exchanges in Cre-LoxP Recombination and Its Basis Suggested by the Crystal Structure of a Cre-LoxP Holliday Junction Complex. *J. Mol. Biol.* **2002**, *319* (1), 107–127. https://doi.org/10.1016/S0022-2836(02)00246-2.
(36) Martin, S. S.; Chu, V. C.; Baldwin, E. P. Modulation of the Active Complex Assembly and Turnover Rate by Protein-DNA Interactions in Cre-LoxP Recombination. *Biochemistry* **2003**, *42* (22), 6814–6826. https://doi.org/10.1021/bi0272306.
(37) Gopaul, Deshmukh; Guo, F.; Van Duyne, G. D. Structure of the Holliday Junction Intermediate in Cre–loxP Site-Specific Recombination. *EMBO J.* **1998**, *17* (14), 4175–4187. https://doi.org/10.1093/emboj/17.14.4175.

(38) Ghosh, K.; Guo, F.; Van Duyne, G. D. Synapsis of loxP Sites by Cre Recombinase. *J. Biol. Chem.* **2007**, *282* (33), 24004–24016. https://doi.org/10.1074/jbc.M703283200.
(39) Guo, F.; Gopaul, D.; Van Duyne, G. D. Asymmetric DNA Bending in the Cre-loxP Site-Specific Recombination Synapse. *Proc Natl Acad Sci* **1999**, *96* (13), 7143–7148. https://doi.org/10.1073/pnas.96.13.7143.
(40) Kawano, F.; Okazaki, R.; Yazawa, M.; Sato, M. A Photoactivatable Cre–loxP Recombination System for Optogenetic Genome Engineering. *Nat. Chem. Biol.* **2016**, *12* (12), 1059–1064. https://doi.org/10.1038/nchembio.2205.
(41) Guillén-Pingarrón, C.; Guillem-Gloria, P. M.; Soni, A.; Ruiz-Gómez, G.; Augsburg, M.; Buchholz, F.; Anselmi, M.; Pisabarro, M. T. Conformational Dynamics Promotes Disordered Regions from Function-Dispensable to Essential in Evolved Site-Specific DNA Recombinases. *Comput. Struct. Biotechnol. J.* **2022**, *20*, 989–1001. https://doi.org/10.1016/j.csbj.2022.01.010.
(42) Warren, D.; Laxmikanthan, G.; Landy, A. A Chimeric Cre Recombinase with Regulated Directionality. *Proc. Natl. Acad. Sci.* **2008**, *105* (47), 18278–18283. https://doi.org/10.1073/pnas.0809949105.
(43) Soni, A.; Augsburg, M.; Buchholz, F.; Pisabarro, M. T. Nearest-Neighbor Amino Acids of Specificity-Determining Residues Influence the Activity of Engineered Cre-Type Recombinases. *Sci. Rep.* **2020**, *10* (1), 13985. https://doi.org/10.1038/s41598-020-70867-5.
(44) Berman, H. M.; Westbrook, J.; Zukang, F.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **1999**, *28* (1), 235–242. https://doi.org/10.1093/nar/28.1.235.
(45) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38. https://doi.org/10.1016/0263-7855(96)00018-5.
(46) Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; de Beer, T. A. P.; Rempfer, C.; Bordoli, L.; Lepore, R.; Schwede, T. SWISS-MODEL: Homology Modelling of Protein Structures and Complexes. *Nucleic Acids Res.* **2018**, *46*, 296–303. https://doi.org/10.1093/nar/gky427.
(47) Jo, S.; Kkim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: A Web-Based Graphical User Interface for CHARMM. *J. Comput. Chem.* **2008**, *29* (11), 1859–1865. https://doi.org/10.1002/jcc.20945.
(48) Park, S.-J.; Kern, N.; Brown, T.; Lee; Im, W. CHARMM-GUI PDB Manipulator: Variouos PDB Structural Modifications for Biomolecular Modeling and Simulation. *J. Mol. Biol. 435* (14). https://doi.org/10.1016/j.jmb.2023.167995.
(49) Lee, J.; Cheng, X.; Swails, J. M.; Yeom, M. S.; Eastman, P. K.; Lemkul, J. A.; Wei, S.; Buckner, J.; Jeong, J. C.; Qi, Y.; Jo, S.; Pande, V. S.; Case, D. A.; Brooks, C. L.; Mackerell, A. D.; Klauda, J. B.; Im, W. CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput.* **2015**, *12* (1), 405–413. https://doi.org/10.1021/acs.jctc.5b00935.

(50) Gao, Y.; Lee, J.; Shan Smith, I. P.; Lee, H.; Kim, S.; Qi, Y.; Klauda, J. B.; Widmalm, G.; Khalid, S.; Im, W. CHARMM-GUI Supports Hydrogen Mass Repartitioning and Different Protonation States of Phosphates in Lipopolysaccharides. *J. Chem. Inf. Model.* **2021**, *61* (2), 831–839. https://doi.org/10.1021/acs.jcim.0c01360.
(51) Liu, H.; Fu, H.; Shao, X.; Cai, W.; Chipot, C. Accurate Description of Cation−π Interactions in Proteins with a Nonpolarizable Force Field at No Additional Cost. *J. Chem. Theory Comput.* **2020**, *16* (10), 6397–6407. https://doi.org/10.1021/acs.jctc.0c00637.
(52) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmuller, H.; Mackerell, A. D. CHARMM36m: An Improved Force Field for Folded and Intrinsically Disordered Proteins. *Nat. Methods* **2016**, *14*, 71–73. https://doi.org/10.1038/nmeth.4067.
(53) Abraham, M. J.; Murtola, T.; Schulz, R.; Pall, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *Softw. X* **2015**, *1* (2), 19–25. https://doi.org/10.1016/j.softx.2015.06.001.
(54) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1998**, *18* (12), 1463–1472. https://doi.org/10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H.
(55) RP, J.; F, L.; GN, M.; A, P. The PDB_REDO Server for Macromolecular Structure Model Optimization. *Int. Union Crystallogr.* **2014**, *1* (4), 213–220. https://doi.org/10.1107/S2052252514009324.
(56) Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **2007**, *9* (3), 90–95. https://doi.org/10.1109/MCSE.2007.55.
(57) Chothia, C.; Levitt, M.; Richardson, D. Helix to Helix Packing in Proteins. *J. Mol. Biol.* **1981**, *145* (1), 215–250. https://doi.org/10.1016/0022-2836(81)90341-7.
(58) Kabsch, W.; Sander, C. Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-bonded and Geometrical Features. *Biopolymers* **1983**, *22* (12), 2577–2637. https://doi.org/10.1002/bip.360221211.
(59) The UniProt Consortium. UniProt: The Universal Protein Knowledgebase in 2025. *Nucleic Acids Research*. November 18, 2024, pp 1–9.
(60) Hoess, R.; Abremski, K.; Irwin, S.; Kendall, M.; Mack Alison. DNA Specificity of the Cre Recombinase Resides in the 25 kDa Carboxyl Domain of the Protein. *J. Moelcular Biol.* **1990**, *216* (4), 873–882. https://doi.org/10.1016/S0022-2836(99)80007-2.
(61) Lavery, R.; Moakher, M.; Maddocks, J. H.; Petkeviciute, D.; Zakrzewska, K. Conformational Analysis of Nucleic Acids Revisited: Curves+. *Nucleic Acids Res.* **2009**, *37* (17), 5917–5929. https://doi.org/10.1093/nar/gkp608.
(62) Pettersen F., E.; Goddard D., T.; Huang C., C.; Meng C., E.; Couch S., G.; Croll I., T.; Morris H., J.; Ferrin E., T. UCSF ChimeraX: Structure Visualization

for Researchers, Educators, and Developers. *Protein Sci.* **2021**, *30* (1), 70–82. https://doi.org/10.1002/pro.3943.
(63) Laskowski, Roman. A.; Swindells, Mark. B. LigPlot+: Multiple Ligand-Protein Interaction Diagrams for Drug Discovery. *J. Chem. Inf. Model.* **2011**, *51* (10). https://doi.org/10.1021/ci200227u.
(64) Ptaszek, A. L.; Li, J.; Konrat, R.; Platzer, G.; Head-Gordon, T. UCBShift 2.0: Bridging the Gap from Backbone to Side Chain Protein Chemical Shift Prediction for Protein Structures. *J. Am. Chem. Soc.* **2024**, *146* (46), 31733–31745. https://doi.org/10.1021/jacs.4c10474.
(65) Jurrus, E.; Engel, D.; Star, K.; Monson, K.; Brandi, J.; Felberg, L. E.; Brookes, D. H.; Wilson, L.; Chen, J.; Liles, K.; Chun, M.; Li, P.; Gohara, D. W.; Dolinsky, T.; Konecny, R.; Koes, D. R.; Nielsen, J. E.; Head-Gordon, T.; Geng, W.; Krasny, R.; Wei, G.-W.; Holst, M. J.; McCammon, J. A.; Baker, N. A. Improvements to the APBS Biomolecular Solvation Software Suite. *Protein Sci.* **2017**, *27* (1), 112–128. https://doi.org/10.1002/pro.3280.
(66) French Pacheco, L. Caracterización espectroscópica de la interacción proteína-ligando: AtLEA4-5 - iones metálicos, LAOBP - aminoácidos y CreCBD - ácidos nucleicos. phD thesis, Universidad Autónoma del Estado de Morelos, Cuernavaca, Morelos, México, 2020. http://riaa.uaem.mx/handle/20.500.12055/1141.
(67) Nagy, G.; Igaev, M.; Jones, N. C.; Hoffmann, S. V.; Grubmuller, H. SESCA: Predicting Circular Dichroism Spectra from Protein Molecular Structures. *J. Chem. Theory Comput.* **2019**, *15*, 5087–5102. https://doi.org/10.1021/acs.jctc.9b00203.

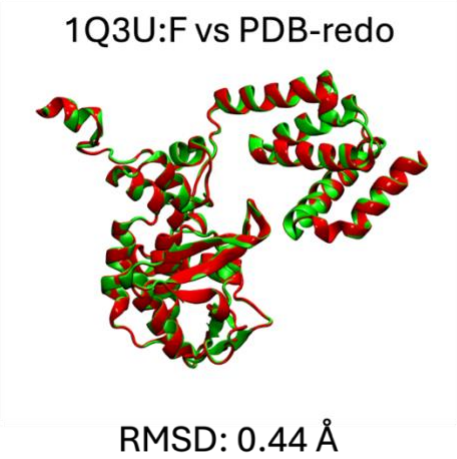
11. Anexos

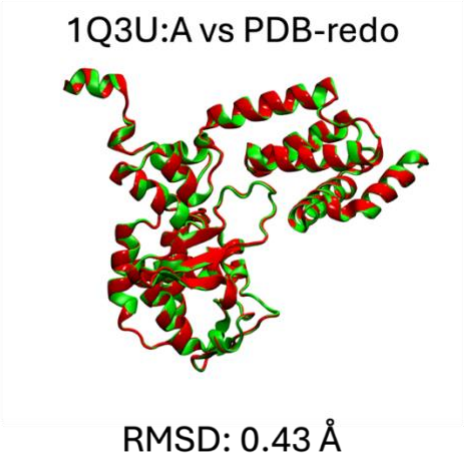
Anexo I.

Publicación asociada a esta tesis. Se puede acceder a través de la siguiente liga: https://pubs.acs.org/doi/full/10.1021/acs.jcim.4c02405

Anexo II.

Comparación estructural entre las estructuras optimizadas y las obtenidas en PDB-redo.

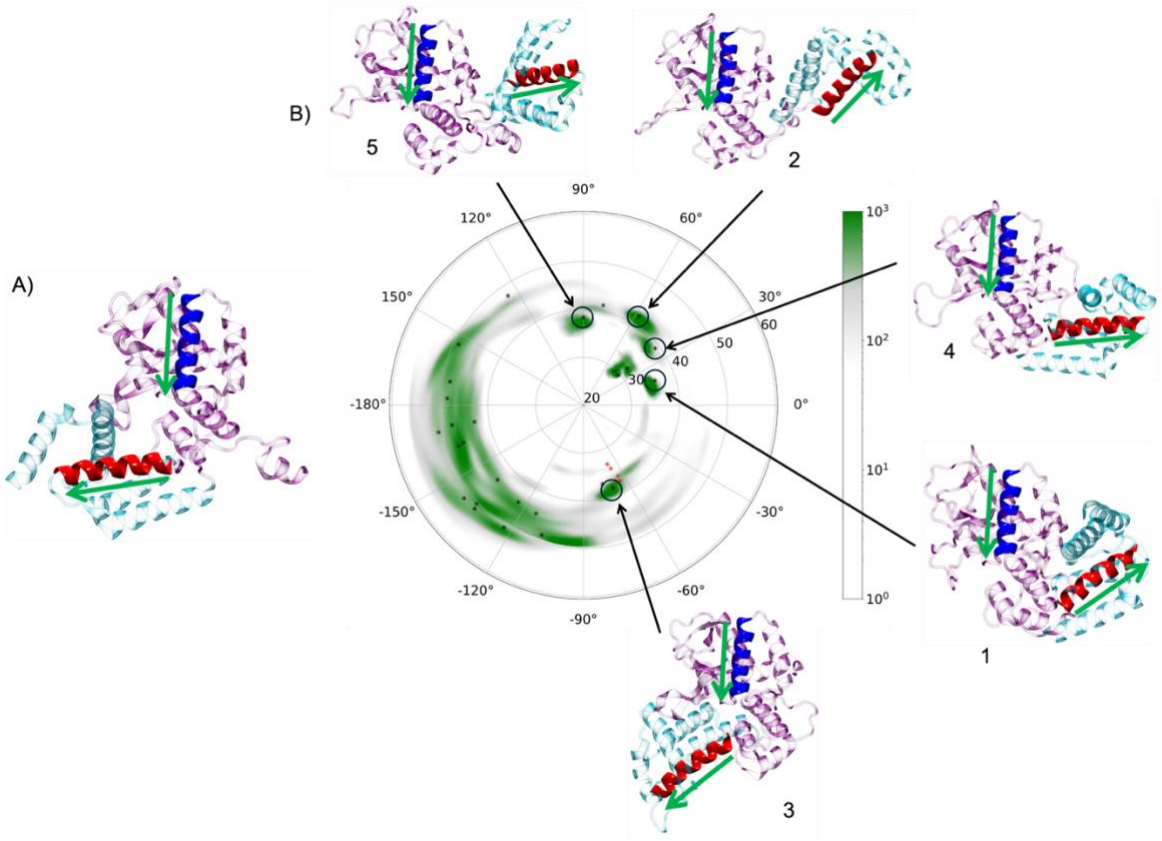




RMSD de la estructura optimizada (verde) contra la obtenida en PDB-redo (rojo), el alineamiento se hizo tomando en cuenta todos los átomos pesados (no hidrógenos), para facilidad de visualización solo se muestra el trazo de cadena principal.

Anexo III.

Estructuras de los cinco grupos más poblados obtenidos por clustering.

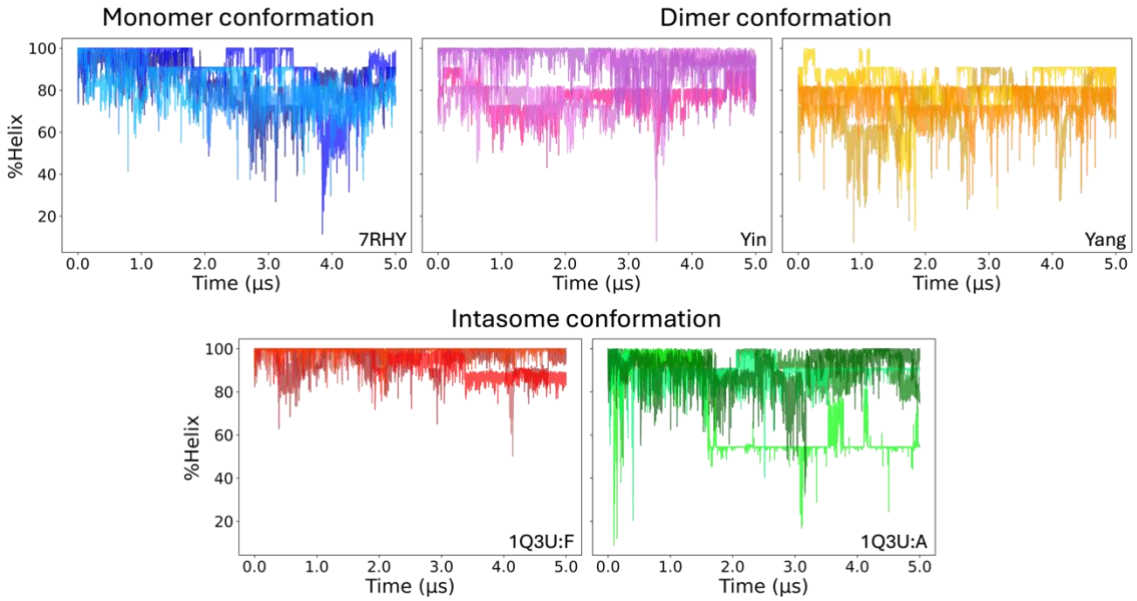


Paisaje conformacional de Cre construido con los ángulos y distancias de las hélices D y G. A) Estructura inicial mostrando la dirección de las hélices de N- a C-terminal. B) Estructuras los cinco clusters más poblados y su localización en el paisaje. El número de miembros en cada cluster está ordenado de 1 (más poblado) a 5 (menos poblado). 1: 4629 miembros, 2: 4006 miembros, 3: 3874 miembros, 4: 3458 miembros, 5: 3049 miembros.

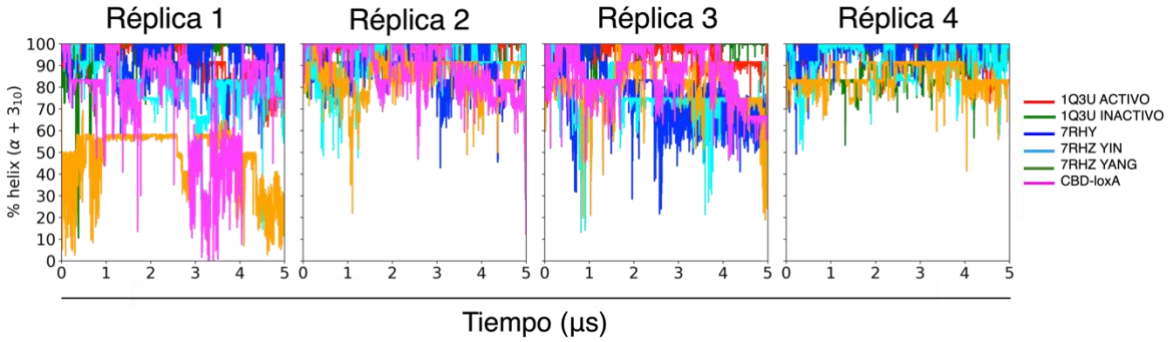
Anexo IV.

Porcentaje de estructura tipo hélice para la hélice A a lo largo del tiempo de simulación para todos los sistemas.

Cre libre:



CBD aislado:



Porcentaje de estructura secundaria tipo hélice para Cre libre y CBD aislado.

Cuernavaca, Mor., a 22 de mayo de 2025

DRA. LINA ANDREA RIVILLAS ACEVEDO
COORDINADORA DEL POSGRADO EN CIENCIAS
P R E S E N T E

Atendiendo a la solicitud para emitir DICTAMEN sobre la revisión de la tesis titulada: **Dinámica molecular de la recombinasa Cre asociada al reconocimiento de *lox*P**, que presenta el M. en C. Marco Antonio Ramírez Martínez (10009558) para obtener el título de Doctorado en Ciencias.

Director de tesis: Dra. Carmen Nina Pastor Colón
Unidad Académica: Instituto de Investigación en Ciencias Básicas y Aplicadas (IICBA)

Nos permitimos informarle que nuestro voto es:

NOMBRE	DICTAMEN	FIRMA
Dr. Carlos Daniel Amero Tello CIQ-UAEM	APROBADO	
Dr. Rodrigo Said Razo Hernández CIDC-UAEM	APROBADO	
Dr. Enrique Rudiño Piñera IBT-UNAM	APROBADO	
Dr. Cesar Millán Pacheco Fac. Farmacia-UAEM	APROBADO	
Dra. Laura Domínguez Dueñas FQ-UNAM	APROBADO	
Dr. Ramón Antonio González García Conde CIDC-UAEM	APROBADO	
Dra. Carmen Nina Pastor Colón CIDC-UAEM	APROBADO	

UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS

Se expide el presente documento con firma electrónica UAEM, soportada por el certificado vigente a la fecha de su elaboración y con efectos plenos de conformidad con los LINEAMIENTOS EN MATERIA DE FIRMA ELECTRÓNICA PARA LA UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS PUBLICADOS en el ÓRGANÓ INFORMATIVO UNIVERSITARIO "ADOLFO MENÉNDEZ SAMARÁ" número 117 de fecha 20 de abril de 2021.

Sello electrónico

CARMEN NINA PASTOR COLON | Fecha:2025-05-22 09:32:18 | FIRMANTE
dq418qpcPWBjpoaGZ86uQPERvPBLvfkUvPpmmS7Fgdg1Sd2WnTNjjK/JIeVno38/DJbiCk2oP3t6YK3VkuvkJibx0VvKfcYCFdAUxXiYabSoxD34UFPnz/TtJy00CjqO9yiBjrR74z0bbDoI4fca8vaSyX/+a71yIN8OAjC8VYsfaBgEifVDREwQ5WmhcR9qZsY5XhxcFiX5IeP3o2AZ+BQNss8MFavXq/VntsUBK7uINa1hvcIPAR/+nFY7gyp+VwCeZpoJjW0q0W03Dwyiro
QPdJD3035R3aOCNKs12v4oD2F92GnZxkQZ9a3p/4N/Rkllq0ct35+od/CN/WX5dA==

RODRIGO SAID RAZO HERNANDEZ | Fecha:2025-05-22 09:39:31 | FIRMANTE
dm4fCP4tY5AE/524ZcYtlrRAZlXe1DN3HRLjBWfi2W1YsoNuQDM35ey6zx1PD3br5X9QmSjo1s+0AwACxc2HSUUFS0aUQY8Fc9Hji2p1NnSBGIZDBubZOUJyEtbCP4f0Rw/7eD
FeHryA9oGd52SZyazY81RnwtjZROE2a9MdD0QtbHqcdBIGTq4zheyQHBJaMtRBQE51m31RmwsDZ7DOlnABEkvVurrXWbtb10oFPsD+cEbpZs8uClb8V1iiiXXJqt1PgiAEmGssp
4iKmmE2X2JJa0tCHV2EdpXbluFTqb316HloFAHOhLzAMQiRNC4Pto7mFslFpH3pjIg31Goc+w==

ENRIQUE RUDIÑO PIÑERA | Fecha:2025-05-22 09:40:02 | FIRMANTE
qSZGVpfzVpG0TcmC5TIHjGccc21wis4CFZNZmy2ucXGiz2WKZJpwjU3HIeMRAiu2Fc/IdVZI7B9uVpnpaexxnF6r/ug3ucHu0d8E2Yt2t0GmOnmmpuE23cAQOQrr5k3EUd26OnZPf
UfyO3ytNm+J98d5BvekRBXyn5W3VaArFgSDjyN1g4ayqGHM2pQh77f58ConuKLL/afl3Lopdh9RnJ5h8V/k1xLGseCVxXSTHMXr549dXURgIPGLH3Zj2+iwKJqxCg3C7qou/xfXZOl
V+gn3l6YzrPVAz2UXlc3f4vdlaDtLih5LlP302ZERiX4E/zV+m8nkS+1EUa/ii2uWVw==

CARLOS DANIEL AMERO TELLO | Fecha:2025-05-22 09:59:45 | FIRMANTE
WtKsONQ9yL7lcxu7KwGnPkNOF8mFEUlzH4TMJxdGk5ZjU7Fu7omXpff9JMkvI3Lf+Hold0717wmUnOd1p8ux4Fc7/sNl+2nWnk0zFbB8HnkILW9/8x0mZ7JHPsEbmLH4JURipaD
78j2ugsplCOSJKp2txRK60qYZHL5tW3lsGZCqEagQDgAF5GZ6dLU8Of4tTOmPfg596uQM295E+Ka4FbIGsSWTp2webQWiCkLOPVxodHK11Y2hUn/Tk75jEw4FdUyjQ13QOpge
4kwSKL1udEJDX4kBFkSg6fPPI+4RhfllQzyRzzd8AjWd+6TqK9s8M8Yeg1p3YgN7wuMv43IsAw==

CESAR MILLAN PACHECO | Fecha:2025-05-22 10:12:04 | FIRMANTE
F4Ii7KZkzVTnWSaGdsAdW7Q+EFUw6KIR+/x7fxiD0myKMiffULqXwXUC1W6aaseEfnMDrVY7iEQFOtjmpJvQH4T2/ft/RYdGr7L24qJ9WhYBAqqAW28Von4JQ8E4sZvM1yww1g
uSUhSiZIqzsiUb9iaF7GWauEeNdv0TluhQBu8ucCzyBAIuFBXxqrKGPfO6onvYBH2i4C4BPSnyJz6d+AdjCwsfMpNh5gJWSvdZY3bcrt/IKwgCpiJYDc9N7f/w5PK/tShPL5Yzxjhe80p
9r7mzA5n4GxvhoHFiy0shrMRAQOC1KV9Toy31jjZBOgIlbylZHlmCyGxGApNmZ0egqA==

RAMON ANTONIO GONZALEZ GARCIA CONDE | Fecha:2025-05-22 12:29:01 | FIRMANTE
EU1c6p+YbGzP3SL+24RuDqF1A7l3N3aHtdLcTZvIiueJE5MunUnu/A/ywcadXntHwAFVIVl5DBqJD7hFpPpn3meSrPhr2On0BfTAfSP/9hGHlIyNJWRaF+oacFrK+lJq4YsoE46NICa
3xPfrbX6iAHiSdz2GUOxvqtlpEFvEWvlpssiVgIdx/IuTCuGiwAnijCvz1rqt4IUG6R99OqPljgVdg2s67xLkWVP7A+hQ2sMj5R4C9R5GzehbFkqD/p2Izu0RAmGtX379U8SHg7pnE39+
yhG+vy7WHSQeItRR1uv8oznvYe8xZua35vjMTtPxJvdh16jQ8VBNoeqGfn59Kg==

LAURA DOMÍNGUEZ DUEÑAS | Fecha:2025-05-22 19:35:22 | FIRMANTE
ulSc2F+AKQNUWD26h8tCoeQfIOsqxXn1yQLOr8ZY0Q64HH4Dj4vPm9QTWgQpaPYJ07V31zZm3X4Mkxf/wJ0H68zvDZtv35S62eGKXbAjUaJJVJuRaPdS8sVV7QcXczXnHVPo
OW1iVfNtSFJUYYQNdA0yH1RzI/KdXp65seCSKPtJGnFoaEbJoi2KFSk5ccvSKXaNo22DdJ229HHix7oskHC/00VdAmj073N5uH356IVilpFZ7qCuzO0rfmP8eWbmdDMVKUtLbnm
9FMjwUQueCPjtEzpJVr7CaeUaLCWfNtiVNN/sTGZrgW86/HYH65mwSSmDTQHmYxT8MDJKAlf6Jw==

Puede verificar la autenticidad del documento en la siguiente dirección electrónica o escaneando el código QR ingresando la siguiente clave:

SNp3V9JXd

https://efirma.uaem.mx/noRepudio/vWY85k2bPTkFA4E4m5qbducfAJYAKioI