



Facultad Politécnica - Universidad Nacional de Asunción

Propiedades conformacionales y energéticas de los Intermediarios de Replicación de DNA circular

por

Victor Manuel Martínez Chamorro

Tesis presentada como requisito para la obtención del Grado de
Doctor en Ciencias de la Computación

en la

Facultad Politécnica - Universidad Nacional de Asunción
Grupo de Investigación en Bioinformática (GBI)

26 de junio de 2020



Facultad Politécnica - Universidad Nacional de Asunción

Propiedades conformacionales y energéticas de los Intermediarios de Replicación de DNA circular

por

Victor Manuel Martínez Chamorro

Orientador: Dra. María José Fernandez De Nestosa

Co-orientador: Dr. Christian Emilio Schearer Serra

Tesis presentada como requisito para la obtención del Grado de
Doctor en Ciencias de la Computación

en la

Facultad Politécnica - Universidad Nacional de Asunción
Grupo de Investigación en Bioinformática (GBI)

26 de junio de 2020



Facultad Politécnica - Universidad Nacional de Asunción

Propiedades conformacionales y energéticas de los Intermediarios de Replicación de DNA circular

Victor Manuel Martínez Chamorro

Aprobado en Marzo de 2020

Prof. Dr. Christian Schearer

Prof. Dr. María José Fernandez De Nestosa

Prof. Dr. Jorge B. Schvartzman

Prof. Dr. Dora B. Krimer

Prof. Dr. Inocencio Ortiz

Prof. Dr. José Colbes

Martínez Chamorro, Víctor Manuel.

Propiedades conformacionales y energéticas de Intermediarios de Replicación de DNA circular / Víctor Manuel Martínez Chamorro

-- San Lorenzo: FP-UNA, 2020.
i - xi; 113 h.: il.; 29 cm.

575.12
M366r

Incluye bibliografías y anexo.

Tesis (Doctorado en Ciencias de la Computación). -- Facultad Politécnica, UNA, 2020

1. Topología del DNA. 2. Bioinformática. 3. Matemática aplicada .I Título.

SCDD: 575.12

M366r

Abstract

Grupo de Investigación en Bioinformática (GBI)

por Victor Manuel Martínez Chamorro

DNA topology changes continuously as replication proceeds. Unwinding of the DNA duplex by helicases is favored by negative supercoiling but it causes the progressive accumulation of positive supercoiling ahead of the replication forks. In covalent closed circular DNA molecules this torsional stress must be removed for the fork to keep advancing. Elimination of this positive torsional stress may be accomplished by topoisomerases acting solely ahead of the fork or simultaneously in the un-replicated and replicated regions after the diffusion of some positive torsional strain from the un-replicated to the replicated regions by the replication forks. In any case, once replication is completed fully replicated molecules are known to be heavily catenated and this catenation derives from pre-catenanes formed during replication. Although there is still controversy as to whether fork swiveling redistributes this positive torsional stress continuously or only as termination approaches, the forces that cause fork rotation and the generation of pre-catenanes are still poorly characterized.

In this thesis, we used numerical simulations, based on the worm-like chain model and the Metropolis Monte Carlo method, to study the interchange of supercoiling and pre-catenation in a naked circular DNA molecule of 4440 bp partially replicated *in vivo* and *in vitro*. We propose that a dynamic gradient of elastic potential energy between the un-replicated and replicated regions drives fork swiveling allowing the interchange of supercoiling and pre-catenation.

Resumen

Grupo de Investigación en Bioinformática (GBI)

por Victor Manuel Martínez Chamorro

La topología del DNA cambia continuamente a medida que avanza la replicación. El desenrollamiento de la doble cadena de DNA, llevado a cabo por la DNA helicasa, provoca la acumulación progresiva de superenrollamiento positivo por delante de las horquillas de replicación. En moléculas de DNA circular covalentemente cerrado este esfuerzo de torsión positivo debe eliminarse para que la horquilla siga avanzando y así continuar la replicación del DNA. La eliminación de este estrés torsional se puede lograr mediante topoisomerasas que actúan por delante de la horquilla o simultáneamente en las regiones no replicada y ya replicada de la molécula después de la migración del estrés torsional positivo hacia la región ya replicada mediante rotación de las horquillas de replicación. De cualquier manera, una vez terminada la replicación, las moléculas completamente replicadas están fuertemente encadenadas entre sí y este encadenamiento deriva de los pre-encadenados formados durante la replicación. Aunque todavía existe controversia sobre si la rotación de las horquillas de replicación redistribuye este estrés torsional positivo de forma continua o sólo cuando se acerca la terminación de la replicación, las fuerzas que causan la rotación de las horquillas de replicación y la generación de pre-encadenados están aún poco caracterizadas.

En la presente tesis utilizamos simulaciones numéricas, basadas en el modelo *worm – like chain*, el método Metrópolis Monte Carlo, para estudiar la distribución del estrés torsional en una molécula desnuda de DNA circular de 4440 pb parcialmente replicada en condiciones *in vivo* e *in vitro*. Proponemos que el gradiente de energía potencial elástica entre ambas regiones del Intermediario de Replicación impulsa el giro de las horquillas de replicación permitiendo el intercambio entre superenrollamiento y pre-encadenamiento.

Agradecimientos

A la Facultad Politécnica de la Universidad Nacional de Asunción (San Lorenzo - Paraguay).

Al Grupo de Investigación de Bioinformática (GBI) y Laboratorio de Bioinformática de la Facultad Politécnica de la Universidad Nacional de Asunción.

Al Laboratorio de Biología Molecular de los Cromosomas del Centro de Investigaciones Biológicas del CSIC (Madrid - España).

A la Dra. María José Fernandez De Nestosa, Dr. Christian Schaerer, Dr. Bernardo Schvartzman y Dra. Dora B. Krimer por el apoyo, orientación me han brindado durante el proceso del desarrollo de esta Tesis.

¡Muchas Gracias!

ÍNDICE GENERAL

Abstract	iv
Resumen	v
Agradecimientos	vi
Lista de Figuras	ix
Lista de Tablas	x
Símbolos y abreviaciones	xi
1 Introducción	1
1.1 Definición del problema	1
1.2 Originalidad de la propuesta	3
1.3 Estructura de la obra	4
1.4 Objetivos	4
1.4.1 Objetivo General	4
1.4.2 Objetivos específicos	4
2 Marco biológico y matemático	6
2.1 Moléculas de DNA en los seres vivos	6
2.1.1 Diferencias entre la estructura y organización del DNA entre organismos procariotas y eucariotas	7
2.1.2 Replicación del DNA	9
2.2 Topología del DNA	10
2.2.1 Índice de ligamiento	12
2.2.2 Índice de torsión	14
2.2.3 Índice de superenrollamiento	16
2.2.4 Teorema de Calugareanu	18
2.3 Superenrollamiento	19
2.4 Encadenamiento	23
2.5 Enzimas encargadas de alterar la topología del DNA: Topoisomerasas	25
2.5.1 Topoisomerasas de tipo I	25
2.5.2 Topoisomerasas de tipo II	25
2.6 Topología de los Intermediarios de Replicación	26
2.6.1 Quiralidad y signo topológico de los Intermediarios de Replicación	29
2.7 Simulación computacional de la topología del DNA	30
2.7.1 Parámetros para la simulación de moléculas circulares de DNA	31
2.7.2 Modelo <i>worm – like chain</i>	32
2.7.3 Selección de número de segmentos y rigidez de doblado	33
2.8 Propiedades termodinámicas de los Intermediarios de Replicación	34

3	Métodos para la simulación de moléculas circulares de DNA	36
3.1	Cálculo numérico de las propiedades topológicas	36
3.1.1	Cálculo del número de cruces de pre-encadenamiento	36
3.1.2	Cálculo numérico del índice de superenrollamiento	38
3.1.3	Cálculo numérico del índice de torsión	39
3.2	Cálculo de la energía elástica	39
3.2.1	Término de energía asociado al diámetro efectivo	40
3.3	Simulación numérica de Intermediarios de Replicación	41
3.3.1	Movimiento de rotación rígida	42
3.3.2	Rotación de las horquillas de replicación	42
3.4	Método para la modificación de la topología de los Intermediarios de Replicación	43
3.5	Procedimiento de simulación numérica	45
4	Resultados	48
4.1	Propiedades termodinámicas en función de las características conformacionales de los Intermediarios de Replicación	48
4.2	Características conformacionales en función de las propiedades termodinámicas de los Intermediarios de Replicación	51
4.2.1	Simulación de Intermediarios de Replicación con un avance del 25% de las horquillas de replicación	51
4.2.2	Simulación de Intermediarios de Replicación con un avance del 50% de las horquillas de replicación	52
4.2.2.1	Intermediarios de Replicación con superenrollamiento negativo	52
4.2.2.2	Intermediarios de Replicación con superenrollamiento positivo	57
4.2.3	Simulación de Intermediarios de Replicación con un avance del 75% de las horquillas de replicación	59
4.2.4	Distribución del estrés torsional en función al tamaño relativo de ambas regiones del Intermediario de Replicación	60
5	Discusión	63
6	Conclusiones	69
	REFERENCIAS BIBLIOGRÁFICAS	71
A	Anexos	79
A.1	Algoritmos de base	79
A.1.1	Polígonos iniciales	79
A.1.2	Cálculo del Writhe	81
A.1.3	Rotación rígida	81
A.1.4	Rotación de dos curvas	82
A.1.5	Verificación del diámetro efectivo	83
A.2	Simulación del superenrollamiento	84
A.3	Simulación del pre-encadenamiento	85
A.4	Cambio de conformación de un RI	87
A.5	Publicaciones originales derivadas de esta tesis doctoral	88

LISTA DE FIGURAS

2.1	Conformaciones de la doble hélice de DNA	7
2.2	Conformación del DNA circular	8
2.3	Replicación semiconservativa del DNA	10
2.4	Micrografía electrónica del DNA superenrollado	11
2.5	Tipos de cruce entre segmentos de una curva	13
2.6	Cinta conformada por dos curvas	15
2.7	Ángulo sólido generado por dos segmentos de una curva	17
2.8	Conformaciones posibles del DNA superenrollado	21
2.9	Nomenclatura de signos topológicos y quiralidad del DNA	22
2.10	Tipos de encadenamiento	24
2.11	Intermediarios de Replicación de moléculas circulares de DNA	28
2.12	Signo topológico y quiralidad de los cruces en los Intermediarios de Replicación	30
3.1	Determinación del número de enlaces mediante suma de signos topológicos	38
3.2	Intermediario de replicación obtenido con el modelo WLC	41
3.3	Movimiento de rotación rígida	42
3.4	Conformación inicial de un Intermediario de Replicación simulado	44
3.5	Procedimiento para modificar la topología de un Intermediario de Replicación	45
4.1	Migración del estrés torsional de la región no replicada a la ya replicada de un RI que alcanzó el 50% de la replicación	50
4.2	Migración del estrés torsional de la región ya replicada a la no replicada de un RI que alcanzó el 50% de la replicación	50
4.3	Distribución del estrés torsional en RIs con un avance del 25% de las horquillas de replicación	52
4.4	Distribución del estrés torsional en RIs que han replicado un 50% de su DNA y acumulan superenrollamiento negativo	54
4.5	RIs que han replicado un 50% de su DNA y presentan una distribución asimétrica del estrés torsional	56
4.6	Distribución del estrés torsional en RIs que han replicado un 50% de su DNA y acumulan superenrollamiento positivo	58
4.7	Distribución del estrés torsional en RIs con un avance del 75% de las horquillas de replicación	60
4.8	Conformaciones iniciales y finales de algunas simulaciones realizadas	61

LISTA DE TABLAS

3.1	Constantes utilizadas en la simulaciones realizadas	47
4.1	Conformaciones y energías iniciales y finales de las simulaciones realizadas	62

Símbolos y abreviaciones

DNA Ácido Desoxirribunucleico

RI Intermediario de Replicación

OC Círculo abierto

CCC Círculo covalentemente cerrado

Lk Índice de ligamiento entre dos curvas

Lk₀ Índice de ligamiento de una molécula circular de DNA relajado

ΔLk Variación del índice de ligamiento de un topoisómero respecto a uno relajado

Tw Índice de torsión de dos curvas entre sí

Tw₀ Índice de torsión de una molécula circular de DNA relajado

ΔTw Variación del índice de torsión de un topoisómero respecto a uno relajado

Wr Índice de superenrollamiento de una curva

RH Dextrógiro

LH Levógiro

Sc Superenrollamiento

Ca Número de encadenados entre dos moléculas circulares de DNA

PreCa Número de pre-encadenados entre dos cromátidas hermanas de DNA

E_b Energía de doblado

E_t Energía de torsión

WLC *worm – like chain*

d_{efectivo} Diámetro efectivo

T Número total de giros de la horquilla de replicación en el procedimiento de simulación

S Número total de pasos en el procedimiento de simulación

E_{NR} Energía de la región no replicada

E_R Energía de la región ya replicada

MD Dinámica molecular

MMC Metrópolis Monte Carlo

para mamá, María Obdulia Chamorro
para papá, Gregorio Martínez Achucarro

Capítulo 1

Introducción

1.1 Definición del problema

Los estudios existentes sobre el comportamiento topológico asociado a la estructura del DNA han brindado importantes conocimientos sobre las propiedades topológicas de las moléculas de DNA plasmídico, tales como el superenrollamiento, encadenamiento y anudamiento. También se tienen importantes conocimientos sobre la topología del DNA cromatínico en organismos eucariotas. Se sabe que estos fenómenos influyen en los procesos biológicos fundamentales en los que participa el DNA de un organismo vivo: replicación, transcripción, reparación y recombinación. Sin embargo, se conoce poco sobre el comportamiento topológico de las moléculas de DNA durante la replicación [1].

La replicación presenta desafíos topológicos derivados de la estructura de doble hélice del DNA, cuando la molécula de DNA se debe desnaturalizar para permitir la copia de las cadenas de polinucleótidos que la componen [2]. Una molécula de DNA que está replicando es denominada Intermediario de Replicación (RI, de *Replication Intermediate*). *In vivo* existe un complejo proteico denominado replisoma que se une a los Intermediarios de Replicación de DNA. Este complejo de proteínas se localiza en las horquillas de replicación, que dividen al RI en dos regiones bien diferenciadas: la región no replicada y la región ya replicada. A su vez, la unión de este complejo proteico a las horquillas de replicación podría establecer restricciones para el paso de energía en forma de estrés torsional entre ambas regiones. Sin embargo, al estudiar estas moléculas aisladas *in vitro* se separa el DNA del complejo proteico encargado de la replicación, posibilitando así la manifestación de procesos espontáneos propios del tránsito de energía entre ambas regiones del RI en forma de torque neto en las horquillas de replicación [3].

Las moléculas de DNA circular se encuentran en varias situaciones *in vivo*; tales como los plásmidos bacterianos, el genoma de un grupo de virus, el DNA mitocondrial o de los cloroplastos vegetales [4] y el DNA circular extracromosómico (eccDNA) presente en *Saccharomyces cerevisiae* y otros eucariotas. Los eccDNAs son capaces de replicar independientemente del resto del genoma y están presentes en una variedad de organismos diferentes. Estas moléculas de DNA circular son de interés por su relación con la clonación de genes [5][6][7]. Los genomas de anellovirus, papilomavirus y poliomavirus consisten en una pequeña cadena de DNA circular, y estos virus son conocidos por su rol en procesos patogénicos. Es sabido que los virus del papiloma humano (VPH) están asociados con un riesgo mucho mayor de cáncer cervical. Los microorganismos responsables del mal de chagas y la leishmaniasis poseen grandes cantidades y formas de DNA circular [8]. Las mutaciones en el DNA mitocondrial contribuyen a algunas enfermedades bien caracterizadas; este DNA también está implicado en varias enfermedades y tumores malignos con etiologías poco aclaradas, y además puede funcionar como una herramienta de diagnóstico. Otros DNA circulares extracromosómicos generalmente se detectan en el cáncer [9][10]. El estudio de la estructura y el modo de formación de estos DNA circulares podrían proporcionar información sobre los mecanismos de inestabilidad del genoma de los mamíferos.

En condiciones fisiológicas, antes de empezar la replicación, el DNA circular tiene almacenado un estrés torsional extra encargado de aliviar el estrés torsional causado por la separación entre las hebras de la doble hélice, durante procesos biológicos fundamentales para las células, tales como la replicación, transcripción y recombinación[11][12]. Este estrés torsional extra almacenado en el DNA se acumula en forma de superenrollamiento, y también debe estar presente en las moléculas hijas una vez terminada la replicación.

Durante la replicación, las enzimas que se mueven a lo largo del DNA, como las DNA y RNA polimerasas, provocan la acumulación de estrés torsional por delante de la dirección de su movimiento. Las enzimas denominadas topoisomerasas evolucionaron en bacterias, arqueobacterias y eucariotas para regular los cambios en la topología del DNA [13]. Las topoisomerasas alivian el estrés torsional modificando la topología del DNA por escisión y re-ligación de las moléculas de DNA. Estas enzimas son el blanco de un número importante de agentes antibacterianos que actualmente están en uso clínico, como las quinolonas y aminocumarinas. Las topoisomerasas son también el blanco de varios de los fármacos más utilizados contra el cáncer, entre los que se encuentran las camptotecinas, que impiden aliviar el estrés torsional dañino para el DNA y se emplean en tumores malignos que son resistentes a otras terapias. Casi todas las formas de cáncer que se consideran curables mediante quimioterapia utilizan inhibidores específicos de algunas topoisomerasas, tales como los etopósidos,

doxorubicina, etc. En conjunto, estas observaciones refuerzan el potencial de topoisomerasas como dianas terapéuticas. Sin embargo, el conocimiento actual sobre la topología del DNA y la manera en que las topoisomerasas regulan los cambios topológicos que tienen lugar durante la replicación, la transcripción, la reparación y la recombinación sigue siendo muy limitado. La comprensión de estos procesos es fundamental a la hora de desarrollar nuevos antibióticos y drogas anticancerígenas basadas en inhibidores específicos de topoisomerasas. Para comprender el mecanismo de acción de estas enzimas durante la replicación del DNA es menester tener conocimiento de las propiedades conformacionales y termodinámicas de las moléculas de DNA.

La mayoría de las técnicas experimentales usadas para el estudio de la topología del DNA, como la electroforesis o la microscopía electrónica, entre otros, alteran las propiedades conformacionales de las moléculas. Esto hace de la comprensión del proceso de replicación del DNA una tarea compleja. Por este motivo, aún queda pendiente la tarea de caracterizar la distribución del estrés mecánico entre ambas regiones de un Intermediario de Replicación y el efecto que tiene la distribución del estrés mecánico en la conformación de la molécula.

Este trabajo muestra un estudio mediante simulación computacional de la distribución de estrés torsional en moléculas de DNA circulares parcialmente replicadas y el efecto que éste tiene en las propiedades termodinámicas de estas moléculas. Por otro lado, también se estudiaron las propiedades conformacionales de los RIs que se encuentran en el estado de equilibrio termodinámico. Con este fin, se modelaron moléculas parcialmente replicadas con las horquillas de replicación detenidas en distintas posiciones antes del punto de terminación de la replicación. Se dotó al modelo de RI con la capacidad de alterar la distribución del estrés mecánico entre sus dos regiones mediante un mecanismo de rotación de las horquillas de replicación, lo cual permitió estudiar el efecto que tienen las deformaciones elásticas de la molécula en las propiedades energéticas del RI, y observar las propiedades conformacionales de las moléculas simuladas.

1.2 Originalidad de la propuesta

La topología de las moléculas circulares de DNA está ampliamente caracterizada tanto a nivel experimental como mediante simulaciones computacionales. Sin embargo, sabemos poco acerca de las propiedades termodinámicas de estas moléculas durante el proceso replicativo, y de cómo estas propiedades influyen en las características conformacionales de las mismas. Es imperioso comprender los efectos que la

replicación causa en las propiedades topológicas del DNA y su relación con el mecanismo de acción de las topoisomerasas durante dicho proceso, así como su influencia en la regulación de la expresión génica. En esta obra se presentan los resultados de simulación computacional de moléculas de DNA circular parcialmente replicadas, siendo este el primer análisis realizado mediante simulación computacional de este tipo de moléculas. En estas simulaciones se estudió el efecto de la distribución del estrés torsional sobre las propiedades termodinámicas de los Intermediarios de Replicación, así como las características conformacionales de estas moléculas cuando alcanzan el estado de equilibrio termodinámico.

1.3 Estructura de la obra

En los siguientes capítulos de esta obra se abordará el siguiente contenido:

En el capítulo dos se presentará una introducción teórica que abarca conceptos biológicos, matemáticos y físicos. En el capítulo tres se describen los conceptos matemáticos y métodos computacionales utilizados para la simulación de los fenómenos topológicos abordados por este estudio. En dicho capítulo también se describe la metodología a seguir para estudiar las propiedades energéticas y conformacionales de los Intermediarios de Replicación. En el cuarto capítulo se muestran los resultados de simulación computacional de las moléculas modeladas. En el quinto capítulo se presenta la discusión de los resultados obtenidos y en el sexto se enumeran las conclusiones. En el anexo se encuentran los algoritmos utilizados para realizar las simulaciones computacionales así como los resúmenes de trabajos presentados en congresos internacionales y los artículos científicos derivados de esta tesis.

1.4 Objetivos

1.4.1 Objetivo General

Estudiar mediante simulación computacional las propiedades conformacionales y termodinámicas de moléculas de DNA circular parcialmente replicadas.

1.4.2 Objetivos específicos

- Desarrollar un modelo de molécula de DNA circular parcialmente replicada.

-
- Estudiar mediante simulación computacional el efecto de las deformaciones causadas por acumulación de estrés torsional sobre las propiedades termodinámicas de moléculas de DNA parcialmente replicadas.
 - Caracterizar las propiedades conformacionales de moléculas de DNA parcialmente replicadas que se encuentran en el estado de equilibrio termodinámico.

Capítulo 2

Marco biológico y matemático

2.1 Moléculas de DNA en los seres vivos

La molécula de DNA (ácido desoxirribunucleico) tiene estructura de doble hélice, constituida por cadenas complementarias y antiparalelas de polinucleótidos [2]. Cada nucleótido en la cadena de polinucleótidos del DNA está constituido por un azúcar pentosa desoxirribosa, un fosfato, y una de cuatro bases nitrogenadas, que incluyen dos purinas, adenina (A) y guanina (G), y dos pirimidinas, timina (T) y citosina (C), donde el grupo fosfato sirve para unir un nucleótido con el siguiente [15]. Los dos polinucleótidos que conforman la doble hélice están unidos entre sí por puentes de hidrogeno entre las bases nitrogenadas (A-T y G-C) que se emparejan de modo complementario, formando un par de bases (pb). El tamaño de una molécula de DNA se mide en términos del número de pb.

En los organismos vivos se encuentran tres tipos de conformaciones distintas del DNA: A-DNA, B-DNA y Z-DNA. La conformación que adopta el DNA depende de su secuencia, el grado y la dirección del enrollamiento que presenta, la presencia de modificaciones químicas en las bases y las condiciones de la solución, tales como la concentración de iones. De las tres conformaciones, la forma "B" es la más común en las condiciones existentes en las células. La doble hélice de B-DNA es una espiral con quiralidad dextrógira, es decir que la doble hélice gira en torno a un eje axial hacia la derecha. Tiene un diámetro de entre 2.2 y 2.6 nm y hace una revolución completa en torno a su eje imaginario cada 10.5 nucleótidos, teniendo esta vuelta completa una longitud de 3.4 nm [16]. La Figura 2.1 muestra una representación de un segmento de una molécula de DNA en sus tres conformaciones posibles. Las otras dos conformaciones de la doble hélice con estructura alternativa difieren en su geometría y dimensiones. La forma "A" es una espiral que tiene una quiralidad dextrógira y es más

ancha que la forma "B". La forma "A" ocurre en condiciones no fisiológicas en formas deshidratadas de DNA. Los segmentos de DNA en los que las bases han sido modificadas por metilación pueden sufrir cambios conformacionales mayores y adoptar la forma "Z" la cual tiene quiralidad levógira, es decir que la doble hélice gira en torno a su eje axial hacia la izquierda [16]. En el resto de esta obra, en cuanto se refiere a la doble hélice de DNA, nos estaremos refiriendo a la forma B-DNA.

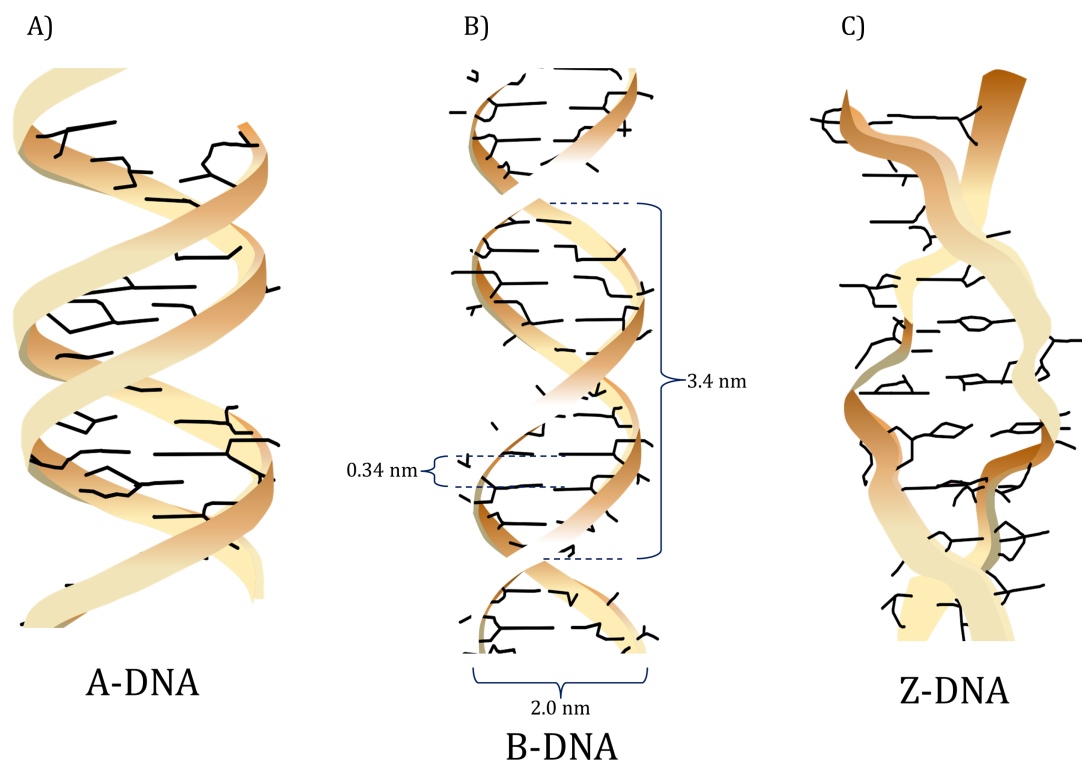


Figura 2.1: Conformaciones de la doble hélice de DNA. A) A-DNA, doble hélice con quiralidad dextrógira que se presenta en condiciones no fisiológicas. B) B-DNA, doble hélice con quiralidad dextrógira, es la conformación más frecuente en la naturaleza. C) Z-DNA, doble hélice con quiralidad levógira.

2.1.1 Diferencias entre la estructura y organización del DNA entre organismos procariotas y eucariotas

Basándonos en la organización de las estructuras celulares, todas las células vivientes pueden ser divididas en dos grandes grupos: Procariotas y Eucariotas. Animales, plantas, hongos y protistas poseen células de tipo eucariota. Sólo las bacterias (eubacterias y arqueobacterias) tienen células de tipo procariota. Se llama procariota a las células sin núcleo celular definido, es decir, cuyo material genético se encuentra disperso en el citoplasma, reunido en una zona denominada nucleóide. En las bacterias se pueden encontrar dos tipos de DNA: el cromosómico y el plasmídico, este último consiste en pequeñas moléculas circulares de DNA extracromosómico que poseen un

mecanismo de replicación independiente. La ventaja de poseer un plásmido es que puede contener genes de resistencia a antibióticos, tolerancia a metales tóxicos, síntesis de enzimas, etc. En las cepas bacterianas, inicialmente estudiadas por microscopía electrónica, se encontró que había un único cromosoma covalentemente cerrado y circular (CCC) [11]. Los plásmidos bacterianos también son moléculas circulares CCC [11]. Un ejemplo de la estructura de las moléculas de DNA CCC se muestra en la Figura 2.2

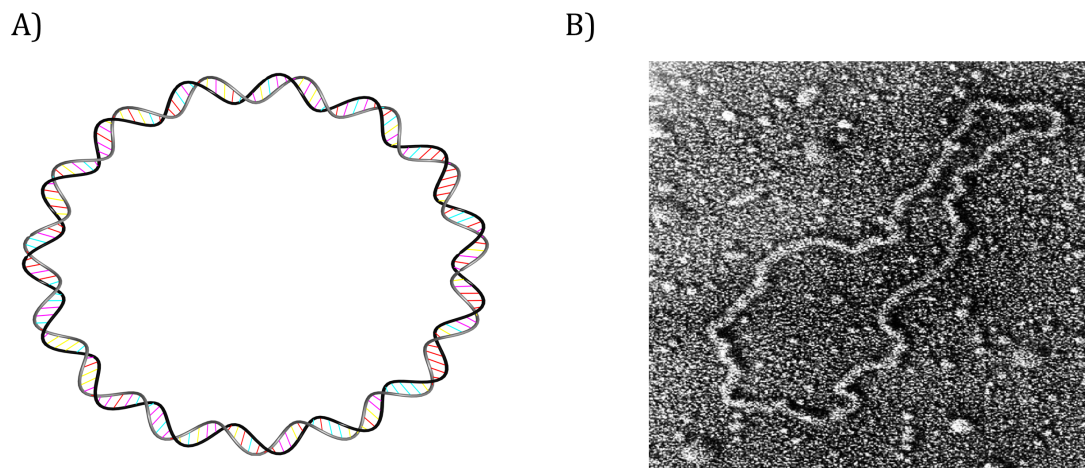


Figura 2.2: Conformación del DNA circular. La doble hélice de DNA, además de estar en forma lineal con ambos extremos libres, puede estar cerrada formando un círculo. A) Representación de una pequeña molécula circular de DNA. B) Micrografía electrónica del plásmido pSC101, una forma de DNA circular. Tomada de Cohen et al. (Cohen, 1973).

En una célula eucariota el material genético suele estar encerrado en un organelo denominado núcleo. Pueden contener varios cromosomas lineales, es decir, con los extremos libres, y el DNA se encuentra plegado en fibras de cromatina complejas. La cromatina es un complejo de DNA e histonas, cuya principal función es la de empaquetar el DNA para que ocupe el menor espacio posible en el interior del núcleo. La unidad estructural de la cromatina es el nucleosoma, el cual está compuesto por un octámero de histonas y DNA. Esta forma de empaquetamiento le da propiedades topológicas al DNA que serán descritas más adelante. Las células eucariotas son más complejas que las procariotas, conteniendo las primeras varios orgánulos especializados para cumplir alguna función vital para la célula. Algunos orgánulos de las células eucariotas también tienen su propio DNA. Dos ejemplos son las mitocondrias (mtDNA) y los cloroplastos de las células vegetales (kDNA), que poseen ambas, al igual que las células procariotas, DNA en forma CCC [17]. El DNA mitocondrial de los parásitos tripanosomátidos, consiste en una red de varios miles de minicírculos (500 - 10000 pb) y unas pocas docenas de maxicírculos (20000–40000 pb) [8][18]. Otra forma en que se presenta el DNA en los organismos eucariotas son los bucles con restricción

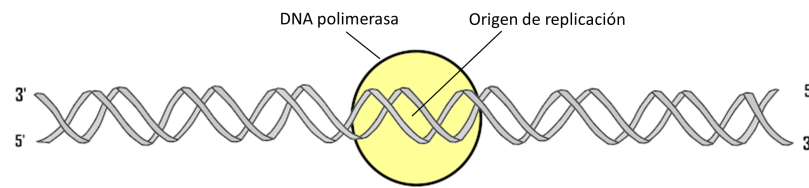
de proteínas, que están involucrados en la regulación de los procesos celulares, incluida la transcripción, la replicación y la recombinación, permitiendo que las regiones distales del DNA entren en contacto y se afecten entre sí [4][19].

2.1.2 Replicación del DNA

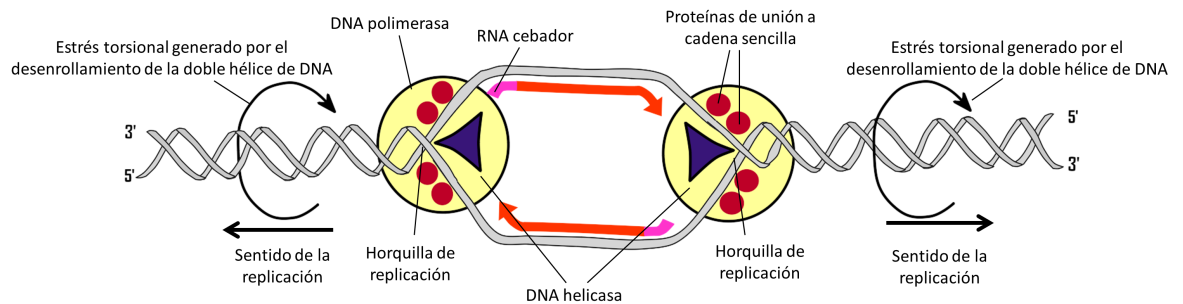
En la replicación semiconservativa, cada cadena de polinucleótidos, emparejados de modo complementario, sirve como molde para sintetizar una nueva cadena [15]. La replicación del DNA se efectúa por la acción de un complejo proteico denominado replisoma. La replicación del DNA empieza en un origen de replicación (Figura 2.3 A). Los puentes de hidrógeno entre las dos cadenas se rompen, permitiendo la apertura de la doble hélice. La enzima responsable de esta separación es la DNA helicasa, creando así una horquilla de replicación que separará la región ya replicada de la no replicada. En el caso de los orígenes de replicación bidireccionales, ambas cadenas se replican al mismo tiempo en dos horquillas de replicación, formando una “burbuja de replicación” (Figura 2.3 B). Cada cadena de la doble hélice puede emparejarse con nuevos nucleótidos complementarios, por la acción de una enzima denominada DNA polimerasa, dando como resultado dos moléculas de DNA, idénticas a la original, cada una formada por una cadena parental y una cadena naciente (Figura 2.3 C).

En las bacterias el DNA circular contiene por lo general un único origen de replicación, y ambas horquillas de replicación avanzan a través de la molécula circular hasta encontrarse en el lado opuesto al origen de replicación, resultando en la formación de dos moléculas circulares de DNA. Los cromosomas eucariotas son moléculas largas de DNA, por lo que contienen varios orígenes de replicación para acelerar el proceso [15]. La posición de las horquillas de replicación se mueve constantemente durante el proceso replicativo.

A)



B)



C)

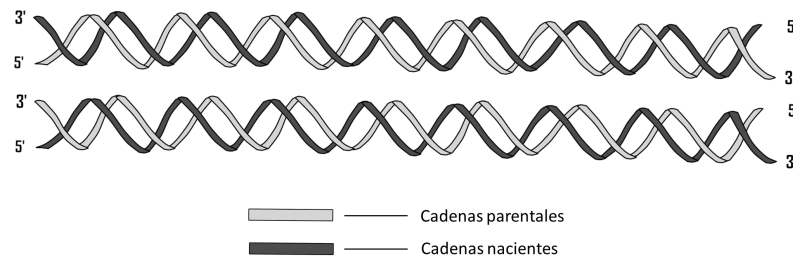


Figura 2.3: Replicación semiconservativa del DNA. A) La síntesis del DNA comienza cuando la DNA helicasa se une a un origen de replicación. B) La DNA helicasa separa las cadenas parentales y las proteínas de unión de cadena sencilla (SSBs, de *Single – Stranded binding*) evitan que las cadenas parentales se unan nuevamente entre sí, mientras que la DNA polimerasa sintetiza las cadenas nacientes, complementarias a las parentales. En el caso de los orígenes bidireccionales, el avance de la DNA helicasa aumenta el tamaño de la burbuja de replicación y en la porción no replicada se acumula el estrés torsional causado por el desenrollamiento de la doble hélice. C) Al terminar la replicación se obtienen dos moléculas hijas, cada una formada por una cadena parental y una recién sintetizada. Modificada de Solomon, 1996 [15].

2.2 Topología del DNA

La topología es una rama de las matemáticas que estudia las propiedades de un cuerpo, relacionadas con su tamaño y forma, que permanecen invariantes cuando éste es plegado, dilatado, contraído o deformado [20]. El descubrimiento de la estructura en doble hélice del DNA llevó inmediatamente a preguntarse cómo se produciría la separación de las hebras entrelazadas para permitir la replicación de la molécula y puso en evidencia el hecho de que durante la replicación surgen problemas topológicos [21][22].

La observación de que en todos los seres vivos el DNA se encuentra superenrollado junto con el descubrimiento de las enzimas encargadas de regular su topología *in vivo*, sentaron las bases para la creación de una nueva disciplina denominada “Topología del DNA” la cual tiene elementos de las matemáticas y la biología molecular [23][24]. Esta nueva disciplina surgió de la necesidad de estudiar el DNA a nivel topológico, siendo los cambios en la topología de la molécula cruciales para el correcto desarrollo de la mayoría de los procesos celulares, tales como la replicación, la transcripción, la reparación o la recombinación, además de afectar el reconocimiento molecular de diversos sustratos conformacionales [25]. Todos estos procesos celulares esenciales están relacionados con las propiedades topológicas de las moléculas de DNA.

La replicación es el fenómeno que provoca más cambios en la topología de las moléculas de DNA. La maquinaria encargada de copiar el DNA abre la doble hélice, induciendo una acumulación de estrés torsional por delante de la horquilla de replicación (Figura 2.3 B).

En 1965, Vinograd y colaboradores observaron al microscopio electrónico que el DNA del virus del polio podía encontrarse tanto en forma lineal como circular y que estos últimos podían presentar distintas formas; relajado o enrollado sobre sí mismo [26]. La apariencia de las moléculas de DNA con superenrollamiento se puede apreciar en la Figura 2.4. A estas moléculas con el mismo tamaño pero distinta forma, asociada a la topología del DNA (más concretamente, de moléculas circulares CCC) se las denomina topoisómeros [26].

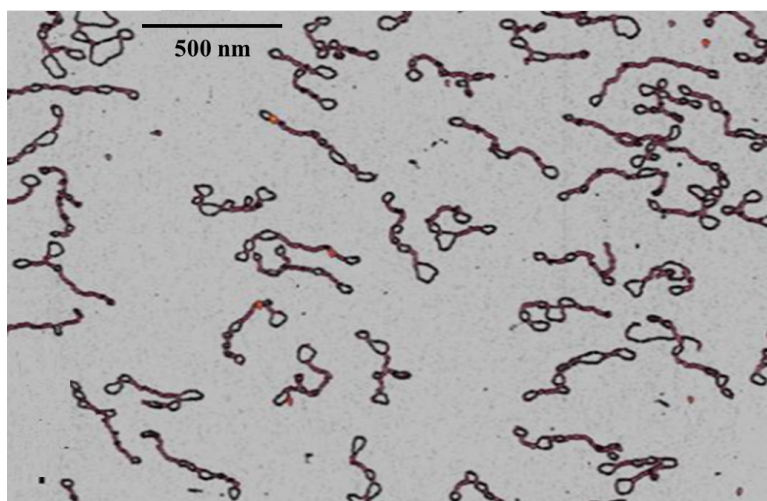


Figura 2.4: Micrografía electrónica del DNA superenrollado. Imágenes obtenidas mediante microscopía de fuerza atómica de plásmidos superenrollados depositados sobre mica. Tomada de Betotti et al. (Betotti, 2018).

Un segmento de DNA en el cual la rotación de sus extremos está restringida se denomina dominio topológico cerrado [27]. La expresión matemática que mejor

describe la topología de un dominio cerrado es el llamado índice de ligamiento o *Linking number*, Lk . El Lk es un invariante topológico que está definido a su vez por dos variables geométricas: el índice de torsión o *Twist*, Tw , y el índice de superenrollamiento o *Writhe*, Wr [11].

2.2.1 Índice de ligamiento

Si una molécula de DNA lineal de doble cadena se cierra en un círculo, formando una estructura covalentemente cerrada, y manteniendo la formación 5'-3' de los enlaces fosfodiéster de cada hebra, el número de vueltas completas de doble hélice de esta nueva molécula será igual al de la molécula lineal original. Este número, que debe ser un entero, corresponde al índice de ligamiento de la molécula de DNA circular cerrada [11]. Para una molécula CCC, con conformación planar, el Lk es la razón entre el número total de pares de bases N y el número de pares de bases por cada vuelta completa de hélice, h , que depende de las condiciones del medio. Un valor en particular para h , se define en condiciones estándar (NaCl 0.2 M, pH 7, 37°) y a menudo se toma como $h_0 = 10.5pb/vuelta$.

Si en la molécula circular no existe ninguna tensión torsional, además de la intrínseca a su estructura helicoidal, se dice que está relajada. En estas condiciones, una molécula de longitud N tendrá una de repetición de doble hélice h_0 , y su índice de ligamiento, al que llamaremos Lk relajado, Lk_0 , tendrá el valor $Lk_0 = N/h_0$.

El índice de ligamiento es una propiedad general de dos curvas cerradas en un espacio tridimensional. Se puede deducir una expresión analítica para el Lk mediante fenómenos de electromagnetismo; suponiendo que tenemos un lazo cerrado, representado por la curva C , que a su vez es descrita por el vector de posición $\mathbf{r}$, que describe cada punto de la curva, en un marco de referencia O , mediante un parámetro t . El lazo transporta una corriente eléctrica I en la dirección en la que la curva C es recorrida por $\mathbf{r}$. Según la Ley de Biot-Savart, el campo magnético en el espacio vacío creado por este lazo cerrado en cualquier punto $\mathbf{r}'$ en el marco de referencia O tiene un valor,

$$\mathbf{B} = \frac{\mu_0 I}{4\pi} \oint_C \frac{(\mathbf{r}' - \mathbf{r}) \times d\mathbf{r}}{\|\mathbf{r}' - \mathbf{r}\|^3} \quad (2.1)$$

donde μ_0 es la permeabilidad del espacio libre [28]. La Ley de Ampere nos dice que la integral del campo magnético a lo largo de una curva cerrada C' , descrita por el vector $\mathbf{r}'$ con parámetro t' , es proporcional a la corriente total que atraviesa el área definida por el lazo cerrado,

$$\oint_{C'} \mathbf{B} \cdot d\mathbf{r}' = \mu_o I_{total} \quad (2.2)$$

Si las curvas C y C' están vinculadas o enlazadas entre sí, entonces la corriente total que pasa a través del lazo C' es igual a un múltiplo entero n de la corriente I , donde este entero es un invariante topológico igual al número de veces que C pasa a través de C' , $I_{total} = nI$ [28].

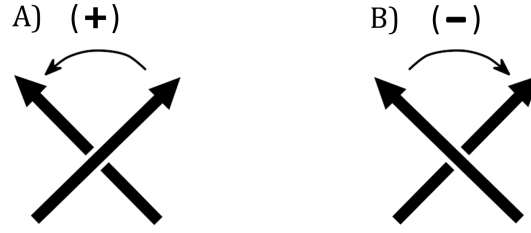


Figura 2.5: Tipos de cruce entre segmentos de una curva. A) Cruce Positivo. El segmento orientado que pasa por arriba debe girar en sentido anti-horario para superponerse con el segmento orientado que pasa por debajo. Para el computo de Lk y Wr este tipo de cruce aporta el valor +1. B) Cruce Negativo. El segmento que pasa por arriba debe girarse en sentido horario para superponerse con el segmento orientado que pasa por debajo. Para el computo del Lk y Wr este tipo de cruce aporta el valor -1.

Introduciendo 2.1 en 2.2

$$\oint_{C'} \left(\frac{\mu_o I}{4\pi} \oint_C \frac{(\mathbf{r}' - \mathbf{r}) \times d\mathbf{r}}{\|\mathbf{r}' - \mathbf{r}\|^3} \right) \cdot d\mathbf{r}' = \mu_o n I \quad (2.3)$$

Reordenando y simplificando nos queda,

$$n = \frac{1}{4\pi} \oint_{C'} \oint_C \left(\frac{(\mathbf{r}' - \mathbf{r}) \times d\mathbf{r}}{\|\mathbf{r}' - \mathbf{r}\|^3} \right) \cdot d\mathbf{r}' \quad (2.4)$$

donde el invariante n es el índice de ligamiento Lk . Esta integral fue desarrollada por Carl Friedrich Gauss, por lo que también recibe el nombre de integral de Gauss [29].

Rigurosamente, el índice de ligamiento también se puede determinar utilizando el número de los cruces de las dos curvas cuando se proyectan en un plano. Esta puede se considerar como una representación bidimensional de las curvas tridimensionales. Cada cruce tiene una dirección y un valor de +1 o -1, dependiendo de la orientación de la curva que pasa arriba con relación a la curva que pasa por debajo (Figura 2.5). El Lk se obtiene sumando los valores asignados a cada cruce y reduciendo el resultado a la mitad.

2.2.2 Índice de torsión

Teniendo en cuenta la estructura de doble hélice de las moléculas de DNA, se puede considerar a sus dos cadenas complementarias como los bordes de una cinta. Asumiendo esto, hay una propiedad denominada índice de torsión, que describe el retorcimiento de esta cinta en torno al eje que pasa a lo largo de la molécula. Se presentan a continuación algunas características propias de las cintas y su retorcimiento tomadas de Klenin et al. [30]:

Dada una curva cerrada C , caracterizada por el vector de posición $\mathbf{a}(t)$, que no tiene intersecciones consigo misma, podemos construir una cinta. Uno de los bordes de esta cinta será la curva C , el otro borde estará constituido por la curva C' , cuyos puntos estarán dados por:

$$\mathbf{b}(t) = \mathbf{a}(t) + \varepsilon \hat{\mathbf{u}}(t) \quad (2.5)$$

donde ε determina el ancho de la cinta y puede ser tan pequeño como se desee. El vector $\hat{\mathbf{u}}(t)$ es un versor normal a la curva C en el punto $\mathbf{a}(t)$.

Si en un sistema de coordenadas ortogonales fijo, xyz , el vector posición $\mathbf{a}$ está dado por el parámetro t como; $\mathbf{a}(t) = x(t)\hat{\mathbf{i}} + y(t)\hat{\mathbf{j}} + z(t)\hat{\mathbf{k}}$, donde las funciones $x(t)$, $y(t)$ y $z(t)$ son continuas y tienen derivada segunda (equivale afirmar que C es una curva suave), entonces podemos definir al vector tangente unitario de $\mathbf{a}(t)$ para cualquier valor del parámetro t como:

$$\mathbf{T} = \frac{\frac{d\mathbf{a}(t)}{dt}}{\left\| \frac{d\mathbf{a}(t)}{dt} \right\|} \quad (2.6)$$

Este vector tiene siempre la misma dirección del vector velocidad, que es sencillamente $\frac{d\mathbf{a}(t)}{dt}$ y donde la rapidez es, por definición, el módulo del vector velocidad [31].

Además de un plano de infinitos vectores normales a $\mathbf{T}$ hay un vector normal que tiene la particularidad de apuntar siempre en la dirección de mayor curvatura de C , dicho vector es conocido como el vector normal principal unitario:

$$\hat{\mathbf{n}} = \frac{\frac{d\mathbf{T}(t)}{dt}}{\left\| \frac{d\mathbf{T}}{dt} \right\|} \quad (2.7)$$

Se define la curvatura en el punto dado por el vector posición $\mathbf{a}(t)$ como:

$$\kappa = \frac{\left\| \frac{d\mathbf{T}}{dt} \right\|}{\left\| \frac{d\mathbf{a}}{dt} \right\|}; \quad (2.8)$$

El vector $\hat{\mathbf{u}}$ normal con el que se genera cada punto de la cinta en la ecuación 2.5 no necesariamente es igual al vector normal principal. Estos conjuntos de vectores se muestran en la cinta de la Figura 2.6.

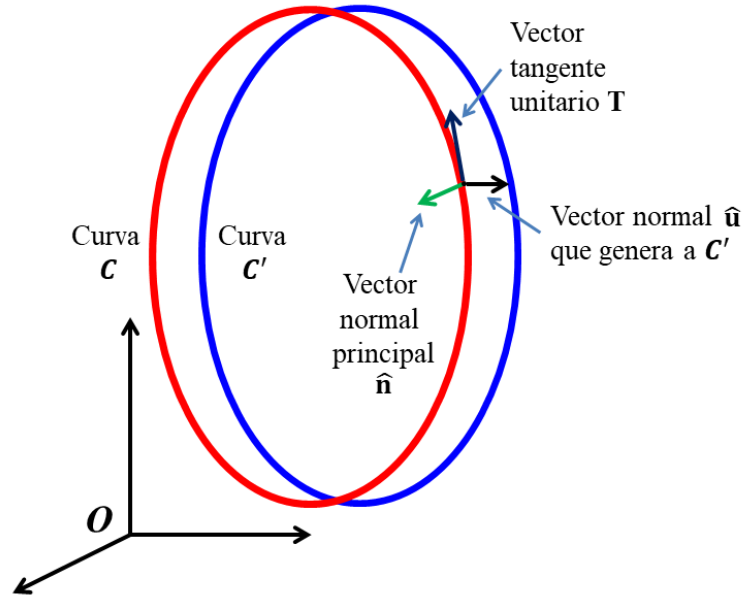


Figura 2.6: Cinta conformada por dos curvas. Cinta creada a partir de las curvas C y C' . Se observa la diferencia entre el vector normal principal, $\hat{\mathbf{n}}$, perteneciente a la curva C y el vector normal generador de la curva C' , $\hat{\mathbf{u}}$. El vector tangente unitario $\mathbf{T}$ es perpendicular a $\hat{\mathbf{n}}$ y $\hat{\mathbf{u}}$ en cada punto de C .

Todas las curvas expuestas hasta ahora pueden ser parametrizadas mediante su propia longitud de arco, s :

$$s = \int_0^t \sqrt{x'(u)^2 + y'(u)^2 + z'(u)^2} du = \int_0^t \left\| \frac{d\mathbf{a}(u)}{du} \right\| du \quad (2.9)$$

Esta parametrización tiene la característica de que la rapidez es constante e igual a 1, es decir, $\left\| \frac{d\mathbf{a}(s)}{ds} \right\| = 1$ [31]. Consecuencia de esta propiedad son las siguientes expresiones de los vectores tangente y normal unitario:

$$\hat{\mathbf{T}} = \frac{d\mathbf{a}(s)}{ds} \quad (2.10)$$

$$\hat{\mathbf{n}} = \frac{1}{\kappa} \frac{d\mathbf{T}(s)}{ds} \quad (2.11)$$

Respecto a la cinta descrita con anterioridad, el valor medio de rotación de la curva C' en torno a la curva C es llamado índice de torsión o *Twist* (Tw). Respecto a la molécula de DNA, el Tw indica el número de veces que las dos cadenas complementarias dan una vuelta completa alrededor del eje axial de la doble hélice. Para calcular el valor del índice de torsión es necesario medir la rotación infinitesimal del vector generador de la curva $\mathbf{b}$ en torno al vector tangente unitario de $\mathbf{a}$, $\mathbf{T}$, para un determinado valor del parámetro s .

La expresión

$$\mathbf{T} \times \hat{\mathbf{u}} = \frac{d\mathbf{a}}{ds} \times \hat{\mathbf{u}} \quad (2.12)$$

proporciona un vector unitario normal a $\hat{\mathbf{T}}$. Al haber una variación en el parámetro ds , el versor $\hat{\mathbf{u}}$ ha variado en $d\mathbf{u}$, apuntando ahora en otra dirección. Si el vector $d\mathbf{u}$ es no nulo entonces siempre tendrá una dirección distinta a la de $\frac{d\mathbf{a}}{ds} \times \hat{\mathbf{u}}$. Una medida de la magnitud de la rotación es entonces

$$\left(\frac{d\mathbf{a}}{ds} \times \hat{\mathbf{u}} \right) \cdot d\mathbf{u} \quad (2.13)$$

y dado que cada rotación es levógira o dextrógira, el signo de la rotación está dado por el signo que tiene este producto escalar [32]. La suma de todas las rotaciones es entonces un promedio ponderado si se divide por la máxima rotación angular de 2π radianes. Así, integrando sobre toda la curva C , el índice de torsión está dado por la expresión

$$Tw = \frac{1}{2\pi} \int_C \left(\frac{d\mathbf{a}}{ds} \times \hat{\mathbf{u}} \right) \cdot d\mathbf{u} \quad (2.14)$$

Para el DNA en estado relajado el Tw es igual al número total de vueltas completas de la doble hélice, en el estado relajado se tiene $Tw_0 = Lk_0$ [4].

2.2.3 Índice de superenrollamiento

El índice de superenrollamiento o *Writhe* de las moléculas de DNA, una de las propiedades topológicas mejor estudiadas, es una medida de la torsión del eje axial de la doble hélice sobre sí mismo. Mide el número de veces que el eje de la doble hélice se pliega sobre sí mismo respecto a un plano o, de modo más intuitivo, es la medida del enrollamiento de la doble hélice sobre sí misma [30].

Para comprender mejor este concepto podemos recurrir al estudio de cómo se comporta el eje de la doble hélice considerándolo como una curva cerrada en el espacio [30]. Para una curva C hay ciertos puntos del espacio, denominados puntos de proyección u observación, desde donde la curva parece cruzarse sobre sí misma. Esto

es, en el cruce de dos segmentos de la curva, el más cercano al punto de proyección parece pasar por encima del que está más alejado.

Hay dos tipos de cruce que se pueden llegar a diferenciar desde cualquier punto de proyección, los positivos (anti-horario) y los negativos (horario) (Figura 2.5). Desde un punto de proyección podemos asignar un valor numérico a un cruce, $+1$ o -1 . Al sumar todos los valores de todos los cruces encontrados en la curva para la dirección de proyección elegida obtenemos un número entero llamado Writhe direccional (Wr_D) [30].

El Wr_D es un entero pero puede no tener el mismo valor para todos los puntos de proyección. Podemos definir el Writhe como el valor promedio del Wr_D desde todos los posibles puntos de proyección [30].

Dados dos segmentos muy pequeños de la curva C , $d\mathbf{r}_1$ y $d\mathbf{r}_2$, que comienzan en los puntos $\mathbf{r}_1$ y $\mathbf{r}_2$ respectivamente, estos segmentos parecen cruzarse entre sí formando un ángulo sólido como el que se muestra en la Figura 2.7. Para conocer este ángulo sólido debemos proyectar el área del paralelogramo formado por $d\mathbf{r}_1$ y $d\mathbf{r}_2$, sobre la esfera de radio $\|\mathbf{r}_1 - \mathbf{r}_2\|$. Esto lo hacemos haciendo la proyección del vector $d\mathbf{S} = d\mathbf{r}_1 \times d\mathbf{r}_2$ sobre el vector unitario:

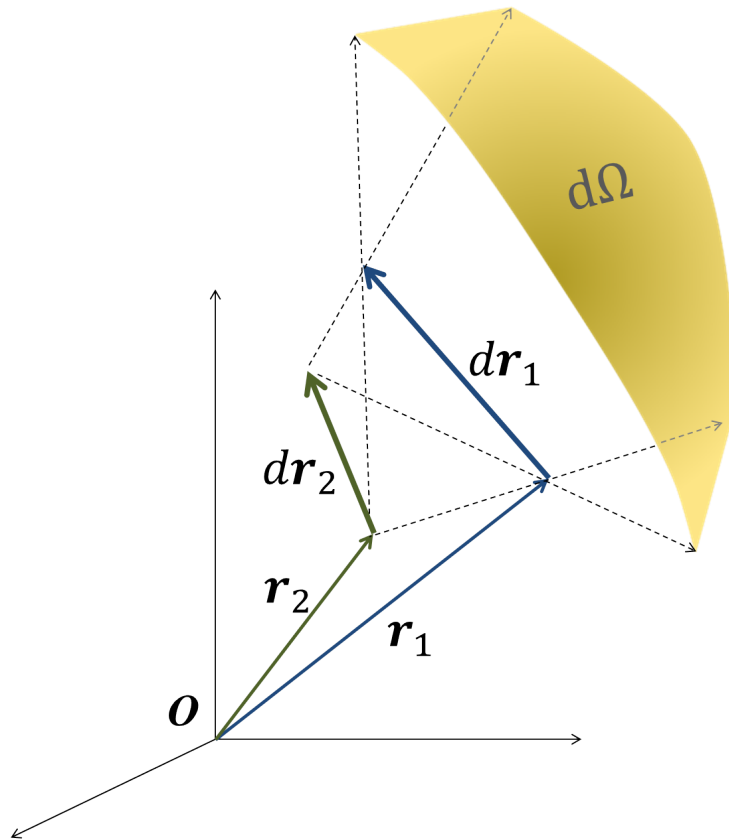


Figura 2.7: Ángulo sólido generado por dos segmentos de una curva. El ángulo sólido $d\Omega$ queda definido desde un punto de proyección desde el cual dos segmentos de una misma curva, $d\mathbf{r}_1$ y $d\mathbf{r}_2$, se cruzan entre sí.

$$\hat{\mathbf{r}} = \frac{(\mathbf{r}_1 - \mathbf{r}_2)}{\|\mathbf{r}_1 - \mathbf{r}_2\|} \quad (2.15)$$

Así, el área sobre la esfera de radio $r = \|\mathbf{r}_1 - \mathbf{r}_2\|$, en la que se ve una intersección entre los segmentos estudiados, está dada por:

$$\hat{\mathbf{r}} \cdot d\mathbf{S} = \frac{(\mathbf{r}_1 - \mathbf{r}_2)}{\|\mathbf{r}_1 - \mathbf{r}_2\|} \cdot (d\mathbf{r}_1 \times d\mathbf{r}_2) \quad (2.16)$$

El ángulo sólido $d\Omega$ se obtiene dividiendo (2.16) por el radio cuadrado r^2 de la esfera, para obtener el área sobre una esfera unitaria:

$$d\Omega = \frac{\hat{\mathbf{r}} \cdot d\mathbf{S}}{r^2} \quad (2.17)$$

o de manera equivalente:

$$d\Omega = \frac{(\mathbf{r}_1 - \mathbf{r}_2) \cdot (d\mathbf{r}_1 \times d\mathbf{r}_2)}{\|\mathbf{r}_1 - \mathbf{r}_2\|^3} \quad (2.18)$$

Cada $d\Omega$ tiene un signo que depende del tipo de cruce y un valor asociado al espacio que ocupa del área de la esfera unitaria. Para conocer el promedio ponderado se debe dividir por el área total de la esfera unitaria, 4π .

Como los segmentos $d\mathbf{r}_1$ y $d\mathbf{r}_2$ pertenecen ambos a la curva C , sus puntos iniciales, $\mathbf{r}_1$ y $\mathbf{r}_2$, deben ambos describir la curva C . Entonces hay que recorrer la curva C dos veces, mediante dos vectores de posición $\mathbf{r}_1$ y $\mathbf{r}_2$ distintos, es decir, construidos con un parámetro distinto. El índice de superenrollamiento Wr está dado por:

$$Wr = \frac{1}{4\pi} \oint_C \oint_C \left(\frac{(\mathbf{r}_1 - \mathbf{r}_2) \times d\mathbf{r}_2}{\|\mathbf{r}_1 - \mathbf{r}_2\|^3} \right) \cdot d\mathbf{r}_1 \quad (2.19)$$

Esta expresión también constituye una integral de Gauss.

2.2.4 Teorema de Calugareanu

Volviendo a considerar la cinta creada por las curvas C y C' , donde C está representada por el vector posición $\mathbf{a}(s)$, y la curva C' es generada por $\mathbf{b}(s) = \mathbf{a}(s) + \varepsilon \hat{\mathbf{u}}(s)$, el teorema de Calugareanu [33][34][35] establece que el índice de ligamiento está dado por:

$$Lk = Wr + Tw; \quad (2.20)$$

donde, el Wr corresponde al índice de superenrollamiento del eje central de la curva (o eje central de la cinta) y Tw es el índice de torsión de sus bordes. El concepto de la cinta toma en cuenta también la estructura de doble hélice de las moléculas de DNA, en donde los bordes de la cinta corresponden a las dos cadenas de polinucleótidos y el eje central de la misma es el punto medio entre estas dos cadenas.

2.3 Superenrollamiento

Las moléculas circulares de DNA covalentemente cerradas o CCCs, constituyen dominios topológicos cerrados y, como hemos visto en el apartado anterior, el Lk se relaciona con el Tw y con el Wr mediante el Teorema de Calugareanu para dominios topológicos cerrados. Si los extremos de la doble hélice de DNA estuvieran libres cualquier tensión estructural interna carecería de efecto ya que las dos cadenas de polinucleótidos podrían rotar sobre sí mismas liberando dicha tensión en los extremos [36]. Por lo tanto, el Wr sólo tiene lugar en moléculas tales como las CCCs o moléculas lineales ancladas a membranas u otros elementos proteicos de tal manera que una región de las mismas, aunque sean lineales, se comportan como un dominio topológico cerrado. Las moléculas de DNA CCCs son típicas del genoma crocomosómico y extracromosómico (plásmidos) de las bacterias, mitocondrias y cloroplastos de organismos eucariotas y también de muchos virus, así como del DNA circular extracromosómico que se encuentra en células humanas, vegetales y animales. Las moléculas de DNA lineales de los cromosomas eucariotas se organizan en bucles con restricciones topológicas mediadas por la unión de proteínas específicas o la interacción con membranas para formar dominios topológicos cerrados equivalentes a los CCCs [12][37].

El efecto que tiene el estrés torsional en las moléculas CCCs se denomina superenrollamiento, y su efecto es el de plegar la doble hélice tridimensionalmente formando superhélices, que tienen una determinada quiralidad. Se dice que en estas condiciones el DNA está superenrollado.

Una forma de caracterizar el superenrollamiento es la variación del índice de ligamiento ΔLk , que se calcula como la diferencia entre el valor que tiene el Lk de la molécula y el que tendría si estuviese relajada, Lk_0 . Cuando una molécula tiene superenrollamiento el estrés torsional también cambia el valor del índice de torsión, por lo cual también hay una variación del índice de torsión ΔTw , que es la diferencia entre el valor de Tw de la molécula y el que tendría si no tuviese estrés torsional extra, Tw_0 .

La variación del índice de superenrollamiento ΔWr es la diferencia entre el valor del Wr presente en una molécula y el valor que tendría si tuviese una conformación plana o relajada, Wr_0 .

El valor promedio del Wr_0 de una molécula CCC en estado relajado es nulo, por lo cual la variación del índice de superenrollamiento se puede expresar como

$$\Delta Wr = Wr \quad (2.21)$$

Se han descrito dos conformaciones para el DNA superenrollado denominadas plectonémica y toroidal. Estas dos formas topológicas se pueden apreciar en la Figura 2.8.

El superenrollamiento plectonómico es un tipo de conformación en la que el eje axial del DNA sigue una trayectoria superhelicoidal alrededor de otro segmento de la misma molécula; dicho de otro modo, la doble hélice del DNA está enrollada alrededor de otra parte de la misma molécula formando una hélice de orden superior [11]. El DNA purificado y en solución presenta un superenrollamiento plectonómico [38]. El superenrollamiento toroidal es un tipo de conformación del DNA superenrollado en donde el eje axial de la molécula sigue una trayectoria superhelicoidal alrededor del eje de un toro imaginario [11]. El DNA de los organismos eucariotas, enrollado alrededor de las histonas en los nucleosomas, presenta un superenrollamiento toroidal que favorece la condensación del DNA necesaria para su empaquetamiento dentro del núcleo [37][38].

Existen además modelos que explican la unión del DNA, en su conformación toroidal, a la superficie de muchas proteínas [11].

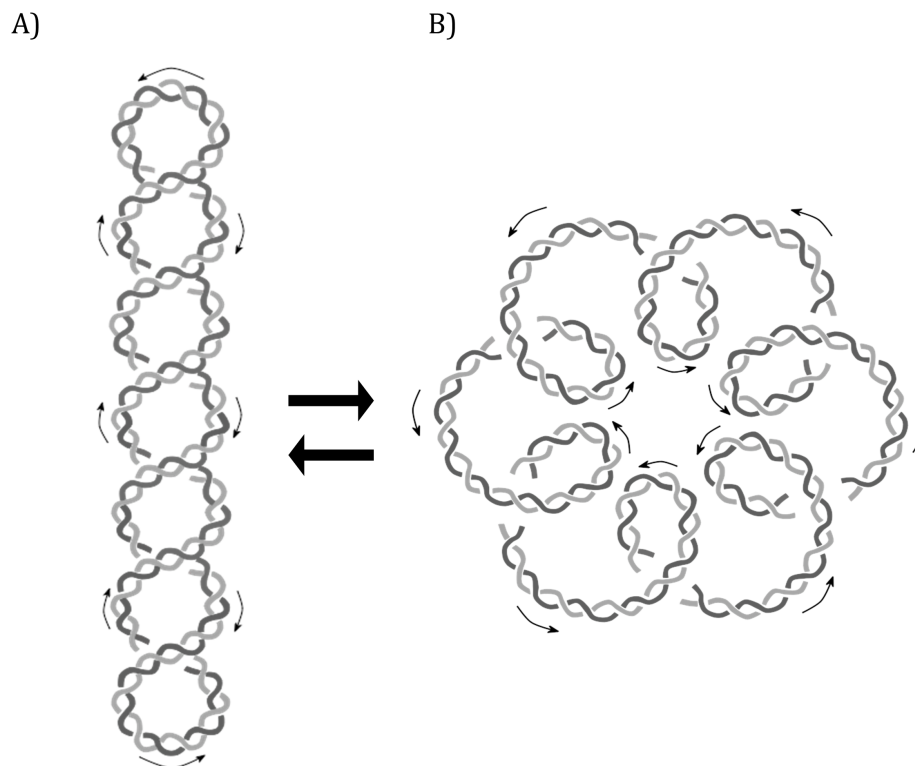


Figura 2.8: Conformaciones posibles del DNA superenrollado. A) Forma plectonémica. La doble hélice de DNA se enrolla sobre sí misma en torno a un eje longitudinal. B) Forma toroidal. El DNA se enrolla de manera que parece plegarse sobre la superficie de un toro imaginario. Modificada de Schwartzman et al. [3].

Hay dos características topológicas del DNA con superenrollamiento que nos interesarán, a saber, el signo topológico y la quiralidad. En la Figura 2.9 A se muestra una manera de diferenciar los signos topológicos tratando a las moléculas de DNA como curvas orientadas. Esto lo hacemos para cruces entre dos segmentos de una doble hélice de DNA. Dicho signo depende del sentido en que debe girar el segmento dirigido que pasa por arriba para superponerse con el que pasa debajo; si se gira en sentido horario el signo es negativo (-) y es positivo (+) en caso contrario. Por otro lado, de acuerdo al sentido en que la doble hélice de DNA está enrollada en torno a un eje imaginario, es decir su quiralidad, la conformación que adopta la molécula puede ser dextrógira (RH de *Right – Handed*), o levógira (LH de *Left – Handed*) (Figura 2.9 B). Por convenio, se dice que el Wr es negativo (-) en el caso de las superhélices dextrógiras y positivo (+) para las levógiras. En la gran mayoría de los seres vivos las moléculas de DNA presentan superenrollamiento negativo, es decir Wr (-) [39][40][41][42][43][44]. Existen algunas excepciones como es el caso de las bacterias termófilas del dominio Archaea, cuyo DNA presenta un Wr (+) que se incrementa con el aumento de la temperatura [45].

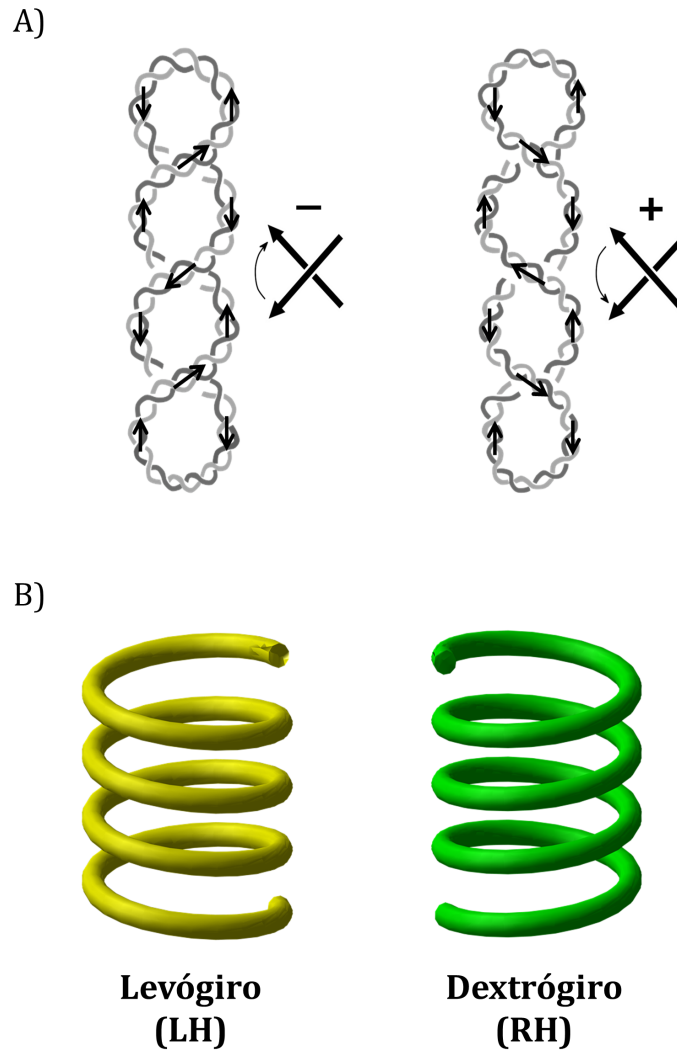


Figura 2.9: Nomenclatura de signos topológicos y quiralidad del DNA. A) El signo topológico se obtiene observando la orientación del segmento de curva que pasa por arriba respecto a la del que pasa por abajo en cada cruce; si el giro para superponer el segmento de arriba con el de abajo es en sentido horario el signo del cruce es negativo (-). El cruce es positivo (+) si el giro es en sentido anti-horario. B) El sentido de giro de una curva en torno al eje central define su quiralidad. Si una curva se eleva girando en sentido horario su quiralidad es levógira (LH), si se eleva girando en sentido anti-horario su quiralidad es dextrógira (RH).

El estrés torsional impuesto al añadir vueltas adicionales a la hélice del DNA se distribuye entre el Tw y el Wr , es decir, ambas variables se compensan entre sí dando como resultado un valor constante de Lk , según la ecuación 2.20. Una consecuencia de esto es que sus variaciones también mantienen constante al ΔLk , es decir $\Delta Lk = \Delta Tw + \Delta Wr$ o, como consecuencia de la ecuación 2.21 $\Delta Lk = \Delta Tw + Wr$.

El Lk puede cambiar cuando se producen roturas en el DNA, las cuales pueden ser de cadena sencilla o de cadena doble y que, en cualquier caso, dan como resultado la eliminación del Wr , ya que permiten que una o ambas cadenas del polinucleótido giren

liberando la tensión en los extremos [46]. Cuando hablamos de cortes de cadena sencilla (*nick*) del DNA *in vitro*, nos referimos a roturas de una de las hebras de la doble hélice que convierten una molécula CCC en una molécula circular abierta y relajada llamada OC (*Open Circle*), esto es, un DNA circular de doble cadena que contiene una rotura en el enlace fosfodiéster en una de las dos cadenas. El Lk también puede cambiar por acción de enzimas especializadas en cambiar la topología de las moléculas de DNA, denominadas topoisomerasas, las cuales serán descritas más adelante.

2.4 Encadenamiento

Una conformación topológica que puede adoptar el DNA es el encadenamiento. Este fenómeno se da cuando dos moléculas de DNA se entrelazan entre sí, una, o más veces. Este tipo de enlazamiento entre distintas moléculas no se puede resolver mediante roturas de cadena sencilla. Este fenómeno topológico fue observado por primera vez en el DNA en 1967 utilizando microscopía electrónica en DNA mitocondrial humano. Más adelante se corroboró que los encadenados son consecuencia de procesos de recombinación y se estudió su formación y resolución *in vitro* [47][48]. Por lo general, los encadenados aparecen como producto final de la replicación de moléculas de DNA.

Existen dos parámetros topológicos para definir los encadenados: el parámetro n , que es el número de cruces que presentan las moléculas encadenadas al proyectarlas sobre el plano; y el número de encadenamiento, Ca , cuyo valor es siempre la mitad de n [49]. Esto nos permite cuantificar el valor Ca para dos curvas que puedan representar moléculas de DNA. Otro método consiste en aplicar la ecuación 2.4 para cuantificar el número de enlace entre ambas curvas, el cual coincide con el valor de Ca .

Los encadenados de DNA formados por dos moléculas, pueden clasificarse de acuerdo a las características topológicas individuales de las moléculas que lo conforman (Figura 2.10). Existen tres tipos:

- Tipo A o CatA: las dos moléculas de DNA circular poseen una rotura de cadena sencilla y, por tanto, están completamente relajadas (ambas son OCs). Un ejemplo de esta forma de encadenamiento puede apreciarse en la Figura 2.10 A.
- Tipo B o CatB: solo una de las moléculas esta relajada (OC), mientras que la otra se encuentra covalentemente cerrada (CCC) y, por tanto, puede albergar superenrollamiento. Un ejemplo de este tipo de encadenamiento se ilustra en la Figura 2.10 B

- Tipo C o CatC: ambas moléculas se encuentran covalentemente cerradas (CCC), pudiendo albergar superenrollamiento [50]. La Figura 2.10 C muestra un ejemplo de este tipo de encadenamiento.

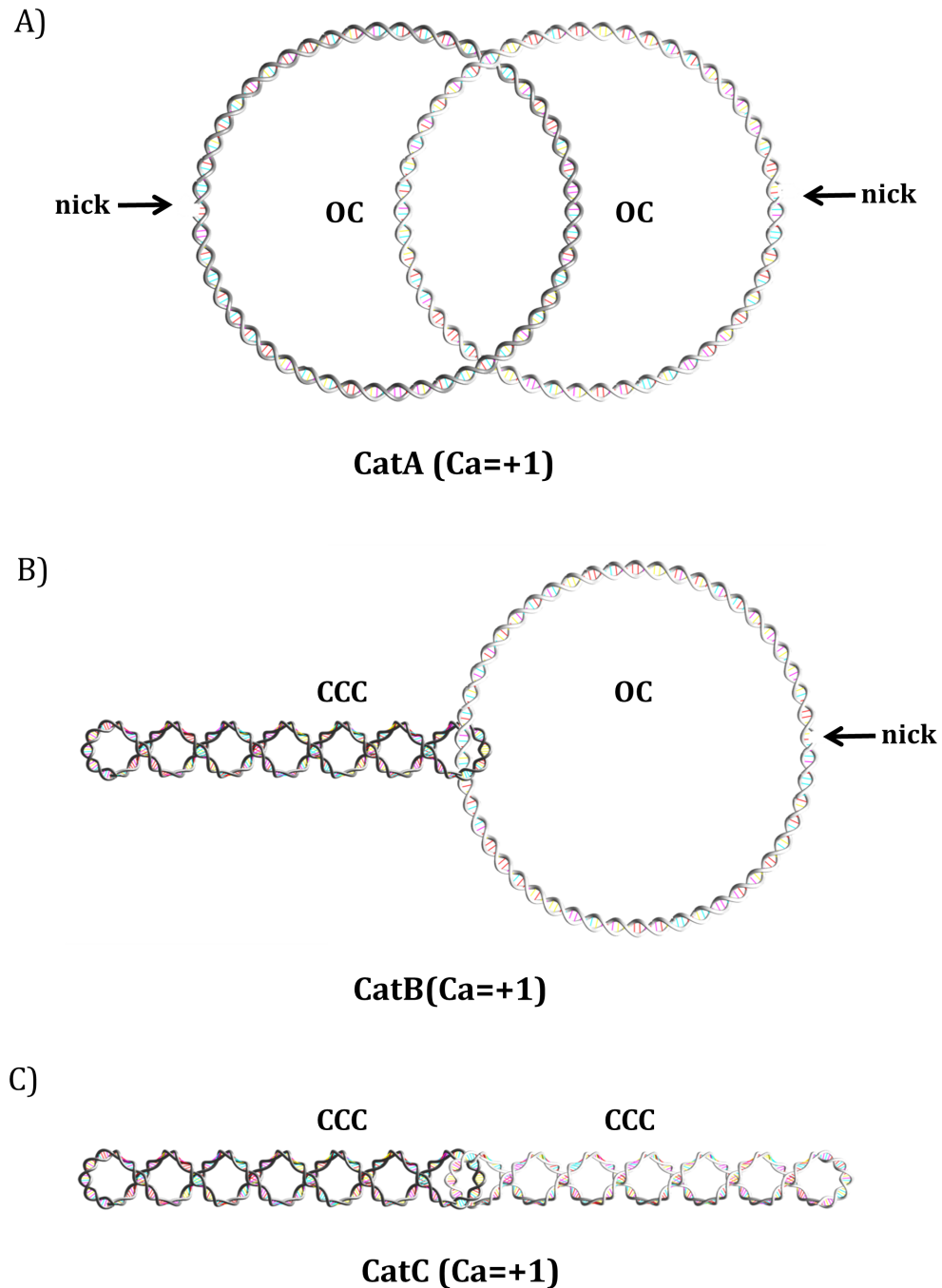


Figura 2.10: Tipos de encadenamiento. Representaciones de moléculas circulares encadenadas con $Ca = +1$. A) CatA, dos moléculas con rotura de cadena sencilla o *nick* (OCs) encadenadas. B) CatB, dos moléculas encadenadas donde una de ellas tiene una rotura de cadena sencilla (OC) y la otra está covalentemente cerrada (CCC). C) Cat C, dos moléculas circulares encadenadas covalentemente cerradas (CCCs). Modificada de Martínez-Robles [49].

2.5 Enzimas encargadas de alterar la topología del DNA: Topoisomerasas

Las topoisomerasas son las enzimas cuya función es modificar la topología del DNA, cambiando el valor del Lk o el Ca . Estas enzimas se encargan de solucionar problemas topológicos generados como consecuencia de procesos celulares tales como la replicación, transcripción, reparación, recombinación y segregación [24]. Por ello, las topoisomerasas son esenciales para el ciclo de vida celular de los seres vivos y se utilizan como dianas de numerosas drogas citotóxicas. Las topoisomerasas de organismos procariotas son dianas de compuestos antibacterianos y antibióticos, y las de organismos eucariotas se usan como dianas terapéuticas en muchos tratamientos antitumorales (Froelich-Ammon and Osheroff, 1995; Kathiravan et al., 2013).

Las topoisomerasas crean roturas transitorias en el DNA que ayudan a resolver los problemas topológicos [24][36][40][48]. Estas roturas transitorias en el DNA van acompañadas de la formación de un enlace fosfodiéster entre un residuo activo de tirosina presente en la enzima y un grupo fosfato de uno de los extremos de la cadena de DNA interrumpida. La topología de la molécula de DNA puede ser modificada durante el tiempo que se mantenga ese enlace. Finalmente, el enlace se rompe y la enzima vuelve a unir los extremos libres del DNA. En base a su mecanismo de acción las topoisomerasas se clasifican en dos grandes grupos, según corten una (tipo I) o ambas hebras (tipo II) de la doble hélice del DNA [46].

2.5.1 Topoisomerasas de tipo I

Las topoisomerasas de tipo I cortan una única cadena de la doble hélice de DNA y la hacen rotar una vuelta completa sobre la otra hebra, luego vuelve a unir los extremos cortados modificando el Lk en valores de ± 1 [51].

Las topoisomerasas de tipo I bacterianas sólo relajan el superenrollamiento (-)[11]. Las topoisomerasas de tipo I en células eucariotas relajan el superenrollamiento (-) y (+)[11].

2.5.2 Topoisomerasas de tipo II

Las topoisomerasas de tipo II cortan ambas hebras de la doble hélice del DNA y transportan la otra doble hélice a través de la que está interrumpida para finalmente volver a unir los extremos cortados [48]. De esta forma hacen que el Lk cambie en valores de ± 2 [47][52]. Todas las enzimas de este tipo son dependientes de ATP [43]. En bacterias existen dos topoisomerasas de tipo II:

- DNA girasa. Es la única topoisomerasa conocida capaz de introducir superenrollamiento (-) en el DNA *in vivo*, utilizando para ello la energía de la hidrólisis del ATP [47].
- Topoisomerasa IV (Topo IV) [54][55]. Es la encargada de eliminar los cruces entre las cromátidas hermanas tanto durante como al final del proceso de replicación del DNA [1][56][57]. Se suele definir como la desencadenasa procariota[42]. Además, es capaz de eliminar ambos tipos de superenrollamiento, sin embargo, es 20 veces más eficiente eliminando los cruces levógiros, presentes en el superenrollamiento (+), que los cruces dextrógiros del superenrollamiento (-) [58][59].

Las topoisomerasas de tipo II eucariota son capaces de relajar ambos tipos de superenrollamiento, pero son incapaces de introducir superenrollamiento (-)[11].

2.6 Topología de los Intermediarios de Replicación

Los plásmidos bacterianos presentan superenrollamiento negativo, es decir, enrollamiento intramolecular de grado superior dextrógiro RH o Wr (-) [60][61]. Este tipo de superenrollamiento es esencial para que la replicación del DNA se pueda llevar a cabo, ya que facilita la separación de las cadenas parentales [25][58][62]. La separación de ambas cadenas permite que sean usadas como molde por la DNA polimerasa durante la replicación. Una vez iniciada la replicación, el Wr (-) también juega un papel importante a la hora de que la elongación progrese correctamente [12].

La replicación se lleva a cabo por acción de un complejo de proteínas denominado replisoma, en el cual la DNA helicasa es la responsable de la apertura de la doble hélice parental [3](Figura 2.3 B). En una molécula circular parcialmente replicada este complejo proteico se encuentra en ambas horquillas de replicación, definiendo dos regiones en la molécula circular replicante, la región no replicada y la región ya replicada. La apertura de la doble hélice llevada a cabo por la DNA helicasa genera estrés torsional que se manifiesta en forma de superenrollamiento (+) en la región no replicada del RI [63][64]. La acumulación de este superenrollamiento (+) podría llegar a detener el avance del replisoma, ya que impediría la separación de las dos cadenas parentales, la cual se encuentra facilitada por el Wr (-) nativo de la molécula. Las DNA topoisomerasas pueden liberar el estrés mecánico acumulado de manera progresiva en las moléculas de DNA. Por lo tanto, en presencia de las topoisomerasas, la replicación puede continuar [12].

Al inicio de la replicación, la región no replicada sería lo suficientemente grande como para que varias de estas topoisomerasas se unan al DNA al mismo tiempo. La acción

de las topoisomerasas podría introducir suficiente Wr (-) para contrarrestar el estrés torsional. Sin embargo, a medida que la replicación progresa, la región no replicada va disminuyendo de tamaño, por lo que cada vez es menor el espacio en el que estas enzimas pueden actuar [1]. La alta procesividad que presenta la DNA helicasa y el rápido avance de las horquillas hace que la actuación de las topoisomerasas no sea suficiente para compensar todo el Wr (+) que se acumula en la región no replicada [1]. Este superenrollamiento (+) normalmente no se puede deshacer debido a las propiedades topológicas de estas moléculas [65][66]. Para resolver este problema, Champoux y Been sugirieron que el superenrollamiento (+) que se acumula transitoriamente por delante de las horquillas de replicación hace girar dichas horquillas, difundiéndose de la región no replicada a la ya replicada [67]. Como consecuencia de la rotación de las horquillas las cromátidas hermanas de la región ya replicada se enlazan entre sí, es decir forman estructuras análogas a encadenados, que se denominaron pre-encadenados (Figura 2.11)[64]. Esta denominación hace referencia a que los pre-encadenados que no son eliminados dan lugar a moléculas encadenadas al finalizar la replicación, que deben ser desencadenadas para garantizar una separación cromosómica adecuada y evitar daños al DNA e inestabilidad cromosómica [68]. Castán et al. demostraron la formación de pre-encadenados durante la replicación en organismos eucariotas y cómo la inhibición de la topoisomerasa 2 conduce a la acumulación de encadenamiento [69].

En el caso de plásmidos bacterianos, en la región ya replicada, la Topo IV, la principal desencadenasa bacteriana, podría actuar eliminando estos pre-encadenados. De esta forma ayudaría a la DNA girasa a compensar la acumulación de superenrollamiento (+) [63][64][70]. En moléculas parcialmente replicadas aisladas *in vitro*, la región ya replicada no puede albergar superenrollamiento debido a que los extremos libres de las cadenas recién sintetizadas pueden girar para liberar torsión, lo que da a estos pre-encadenados propiedades topológicas similares a los encadenados de tipo CatA.

Un estudio en levaduras afirma que la rotación de las horquillas interrumpe intrínsecamente la duplicación normal de cromosomas [71]. Estudios publicados por Cebrián et al. en *E. Coli* demuestran la formación de los pre-encadenados durante la replicación, confirmando el modelo de Champoux y Been [1][72][73][74]. Según este modelo, en un momento específico durante el proceso replicativo, la molécula de DNA tiene cierto nivel de ΔLk . En la región no replicada el ΔLk se acumula en forma de superenrollamiento, mientras que en la región ya replicada lo hace en forma de pre-encadenamiento (Figura 2.11) [70]. Estos cambios topológicos fueron confirmados mediante microscopía electrónica, Peters et al. observaron que los Intermediarios de Replicación purificados adoptan una estructura con superenrollamiento plectonómico en la región no replicada y pre-encadenamiento en la región ya replicada (Figura

2.11D). Además, sugieren que la distribución ΔLk depende del cociente entre las longitudes de las regiones no replicada y ya replicada de la molécula [64][70].

Basándonos en el modelo propuesto por Champoux y Been podemos afirmar que: (1) la energía proveniente del estrés torsional puede difundirse de una región a otra del RI y (2) las topoisomerasas pueden actuar en ambas regiones del RI.

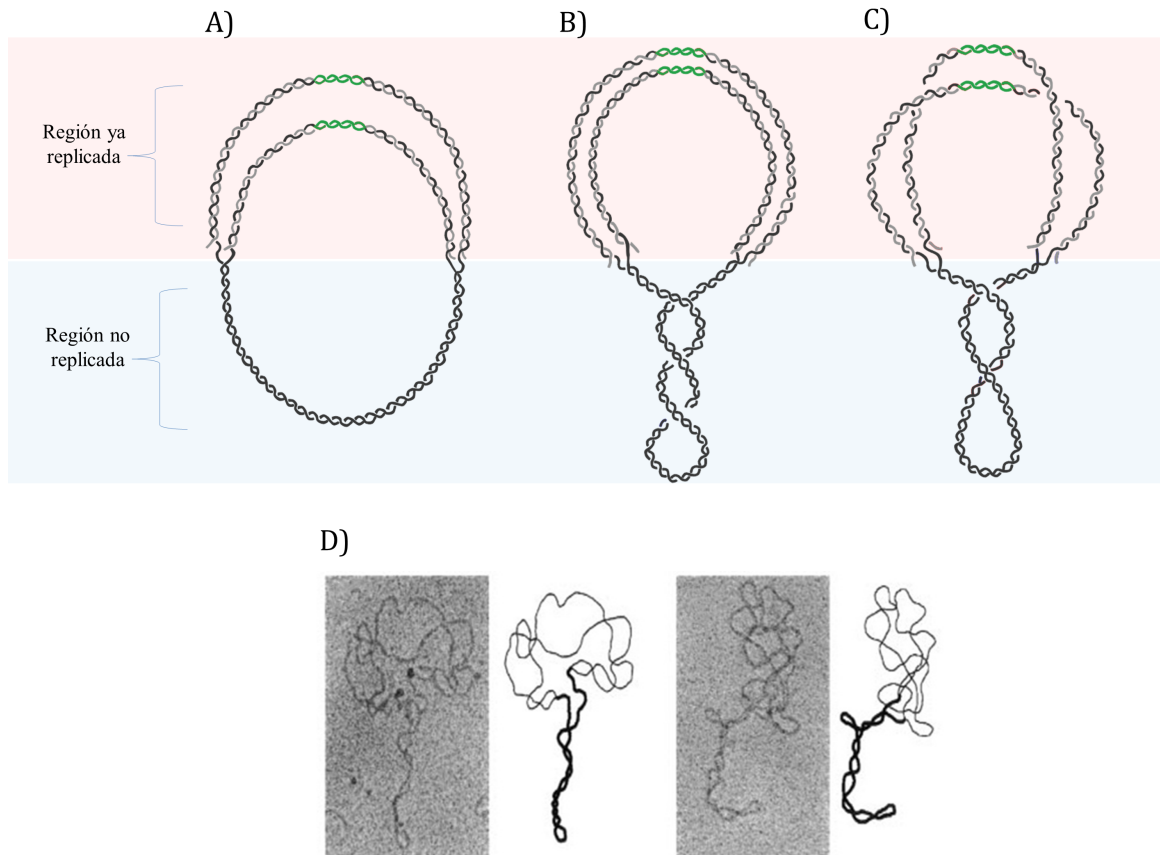


Figura 2.11: Intermediarios de Replicación de moléculas circulares de DNA. Representación de las regiones no replicada y ya replicada. Las cadenas parentales se ilustran en negro, las cadenas recién sintetizadas de color gris y el origen de replicación en verde. A) RI completamente relajado, con $\Delta Lk = 0$. B) RI con estrés torsional, todo el ΔLk se encuentra confinado en la región no replicada en forma de superenrollamiento plectonómico. C) Distribución del estrés torsional en un RI según el modelo de Champoux y Been. El ΔLk se redistribuye entre la región no replicada, en forma de superenrollamiento, y la región ya replicada, en forma de pre-encadenamiento entre las doble hélices recién replicadas. D) Micrografías electrónicas de RIs con ΔLk distribuido entre ambas regiones de la molécula. Modificada de Peter et al. [64].

2.6.1 Quiralidad y signo topológico de los Intermediarios de Replicación

En un RI con las horquillas de replicación detenidas, en cual el estrés torsional no puede pasar libremente de la región no replicada a la ya replicada, la región no replicada presentará superenrollamiento RH (-), y la región ya replicada no mostrará ningún pre-encadenado (Figura 2.12 A). En condiciones *in vitro*, después de la desproteización y la consecuente eliminación de replisoma, el estrés torsional acumulado se distribuirá entre ambas regiones mediante la rotación de las horquillas de replicación. La doble hélice de la región no replicada perderá parte de su superenrollamiento RH (-) mientras que las cadenas de la región ya replicada se enrollarán entre sí en sentido levógiro LH (-) (Figura 2.12 B) [75]. Lo contrario ocurre en el caso del superenrollamiento positivo, LH (+). Resumiendo, al pasar el superenrollamiento de la región no replicada a la ya replicada lo hace con quiralidad opuesta pero conservando el signo topológico [3]. En el caso de RIs obtenidos *in vitro*, éstos presentan cruces RH (-) en la región no replicada debido al superenrollamiento negativo y cruces LH (-) en los pre-encadenados de la región ya replicada [1]. La Figura 2.12B se muestra un modelo de Intermediario de Replicación *in vitro*. Esto sería posible si, antes de la desproteización, el superenrollamiento RH (-) tuviera un valor absoluto mayor que el superenrollamiento LH (+) causado por el avance el replisoma.

En la Figura 2.12 C se muestra un modelo de Intermediario de Replicación *in vivo* propuesto por Jorge B. Schvartzman et al., en el cual las cadenas nacientes de la región ya replicada pueden girar una alrededor de la otra, introduciendo pre-encadenamiento [3]. La migración del estrés torsional, en forma de superenrollamiento LH (+), desde la región no replicada causa la formación de pre-encadenados con cruces RH (+) en la región ya replicada, los cuales serán eliminados por la Topo IV. Al terminar la replicación los cruces pre-encadenamiento que no fueron eliminados por la Topo IV se transforman en encadenamiento entre las cromátidas hermanas.

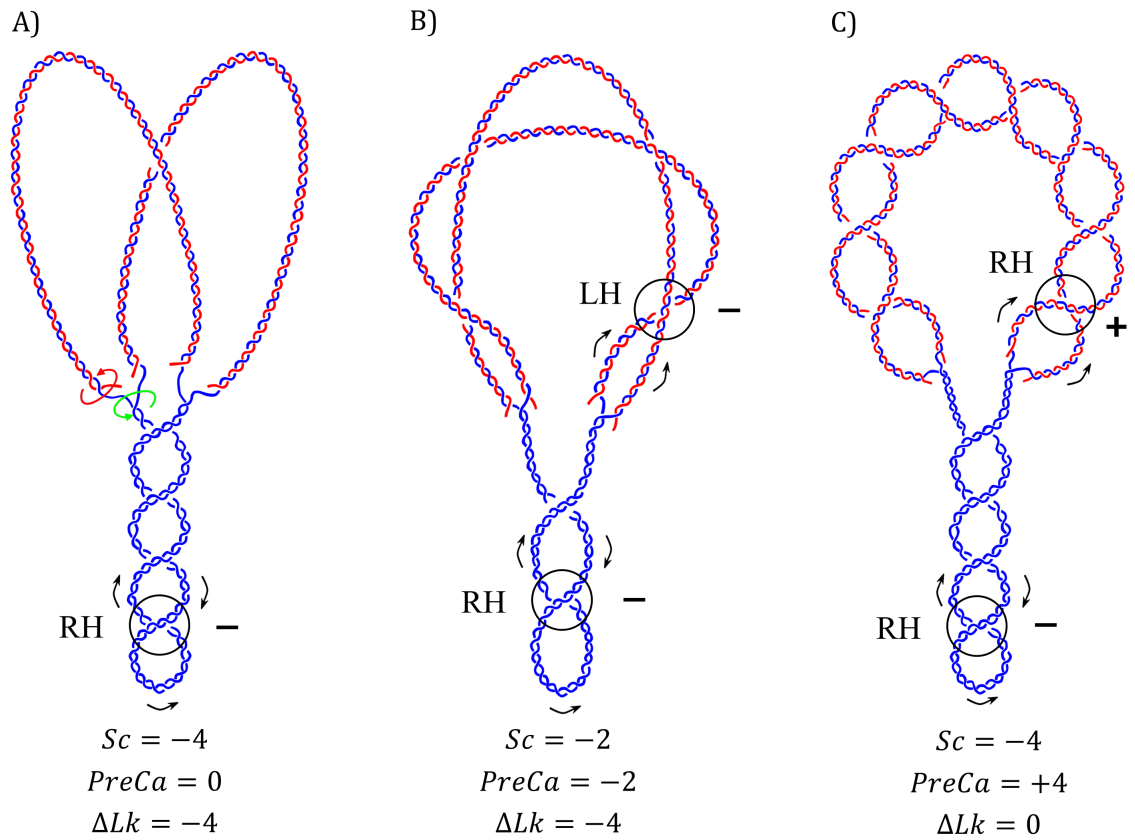


Figura 2.12: Signo topológico y quiralidad de los cruces en los Intermediarios de Replicación. A) Intermediario de Replicación con todo el estrés torsional acumulado en la región no replicada en forma de superenrollamiento con cruces RH (-). Las horquillas de replicación pueden girar libremente (flecha verde), mientras que la región ya replicada no puede albergar superenrollamiento debido a que los extremos libres de las cadenas recién sintetizadas pueden girar para liberar estrés torsional (flecha roja). B) RI en condiciones *in vitro*. El estrés torsional acumulado en forma de superenrollamiento RH (-) se distribuye entre ambas regiones formando pre-encadenados con cruces LH (-). C) RI en condiciones *in vivo*. La región no replicada mantiene un superenrollamiento RH (-) neto mientras que la región ya replicada alberga pre-encadenados RH (+), resultantes de la difusión del exceso de superenrollamiento LH (+) generado por la acción de DNA helicasa. Modificada de Cebrián et al. [1].

2.7 Simulación computacional de la topología del DNA

Las simulaciones computacionales de conformaciones recurren a métodos simplificados de la estructura de las moléculas de DNA, debido a la gran complejidad de las mismas. Los métodos de dinámica molecular (MD de *Molecular Dynamics*) se implementan con tiempo computacional muy alto incluso para pequeños sistemas de moléculas [76]. Por esta razón, actualmente la mayoría de las simulaciones de las conformaciones de DNA se llevan a cabo por el método de Metrópolis Monte Carlo (MC). El método de Metrópolis MC simula propiedades de un sistema utilizando un procedimiento que implica números aleatorios y probabilidad con el fin de encontrar mínimos locales de energía. En el sentido

más específico de simulaciones moleculares, se desarrollan modelos de conformación de una molécula de DNA que se pueden utilizar para calcular energía y a continuación se provocan cambios aleatorios en la conformación para comparar las energías de un conjunto de conformaciones, ya sea con el fin de encontrar la conformación más estable o para generar un conjunto de conformaciones posibles. El método específico usado en la mayoría de los estudios de DNA, así como en muchos otros, fue desarrollado por Metrópolis y colaboradores en 1953 [77].

A partir de una conformación inicial se calcula su energía inicial E_1 , luego, se induce un cambio conformacional aleatorio en el modelo para generar una nueva conformación de prueba y se calcula la energía de esta conformación de prueba, E_2 . La conformación de prueba es aceptada inmediatamente si tiene una energía más baja que la del predecesor ($E_2 < E_1$), pero si la energía es mayor, la conformación se acepta al azar, con una probabilidad p que depende exponencialmente de la diferencia de energía entre las dos conformaciones:

$$p = e^{\frac{(E_1 - E_2)}{k_b T}}; \quad (2.22)$$

Esta fórmula representa la probabilidad de que una conformación determinada con $E_1 < E_2$ sea aceptada bajo las condiciones térmicas indicadas. Si se acepta la conformación de prueba, se utiliza como la base para la generación de una nueva conformación. Si se rechaza, la conformación anterior se utiliza para el siguiente paso. De esta manera, después de muchas conformaciones de prueba, típicamente cientos de miles, las simulaciones convergen a conformaciones de una distribución térmica de las moléculas estables. Estos métodos fueron aplicados por primera vez para una conformación circular de DNA por Le Bret y Vologodskii et al. [78][79][80][81].

2.7.1 Parámetros para la simulación de moléculas circulares de DNA

Para la simulación de las propiedades termodinámicas y conformacionales de moléculas circulares de DNA se deben tener en cuenta ciertas propiedades intrínsecas del tipo de moléculas que se quiere representar; la carga eléctrica de la molécula, su rigidez ante los esfuerzos de doblado, torsión y estiramiento, diámetro, entre otros.

Dos propiedades muy importantes para la simulación son el diámetro de la molécula y la rigidez en función de su longitud. Debido a la carga eléctrica negativa del DNA, dos segmentos separados se repelen entre sí. La fuerza repulsiva depende de la concentración de iones del medio en el que se encuentra la molécula. A la distancia mínima a la que pueden estar dos segmentos distintos de moléculas de DNA, a causa de la repulsión

eléctrica, se le llama diámetro efectivo y corresponde al mismo diámetro de la doble hélice [11]. El diámetro efectivo y la longitud de la molécula definen un volumen denominado volumen excluido. La energía necesaria para doblar la molécula en un círculo es menor mientras mayor es la longitud total de esta. La flexibilidad de la molécula de DNA se describe normalmente por un parámetro llamado longitud de persistencia, a , que describe la escala de longitud sobre la cual la rigidez de la hélice de DNA es importante [11][80]. Se puede definir la longitud de persistencia como la longitud sobre la que existe una cierta correlación entre la dirección inicial y final de la hélice. Después de una distancia de una longitud de persistencia a , debido al movimiento térmico, la desviación media del eje de la hélice a partir de su dirección inicial será un radián, y después de dos longitudes de persistencia, $2a$, ya no habrá correlación entre las direcciones inicial y final, con un ángulo medio de deflexión cercano a 90° . El valor $2a$ también se conoce como la longitud estadística de Kuhn [11][80]. El valor de la longitud de persistencia está directamente relacionado con la rigidez de doblado de la molécula y con la temperatura absoluta del medio. Otra característica importante de una conformación simulada de superenrollamiento es el ΔLk , el cual se había definido anteriormente como la diferencia entre el Lk de la molécula y el Lk_0 de la molécula si esta estuviese relajada.

2.7.2 Modelo *worm – like chain*

El modelo que se utiliza en esta obra para simular moléculas de DNA es conocido como *worm – like chain* (WLC). En este modelo las moléculas están representadas por una serie de segmentos rectos de longitud constante [80][81][82]. Si se asigna a cada segmento de la curva una distancia mínima a la que se le permite estar de otro segmento, entonces se puede suponer que la curva es ahora un cilindro o tubo impenetrable. Para emular una molécula circular de DNA la curva de segmentos rectos debe ser cerrada y con un número finito de segmentos. Además, el método de Metrópolis MC demanda adquirir un valor de energía para cada conformación de prueba creada.

La energía potencial elástica de las moléculas simuladas en esta obra tendrá dos componentes: una energía elástica de doblado, E_b , y una energía elástica de torsión, E_t . La energía de doblado está asociada al ángulo que existe entre dos segmentos de la curva. Para el modelo WLC la energía de doblado de la molécula está dada por:

$$E_b = k_b T \alpha \sum_{i=1}^N \theta_{i,i+1}^2 \quad (2.23)$$

donde N es el número de segmentos que componen la molécula modelada, la sumatoria se extiende a todas las uniones entre los segmentos rígidos, k_b es la constante de Boltzmann, T es la temperatura absoluta (297 K), $\theta_{i,i+1}$ es el desplazamiento angular (en radianes) del segmento i relativo al segmento $i + 1$ y α es la constante de rigidez de doblado [80].

La energía de torsión, E_t , es una función cuadrática de la variación del Tw de la cadena respecto a la configuración con Lk_0 , denominada ΔTw . No se computa ΔTw directamente porque el Tw es una propiedad de la estructura de doble hélice del DNA que no está incluida en el modelo WLC. Por este motivo, en este modelo el ΔTw se calcula utilizando el Teorema de Calugareanu:

$$\Delta Tw = \Delta Lk - \Delta Wr \quad (2.24)$$

Así la energía elástica de torsión puede ser expresada como:

$$E_t = \frac{2\pi^2 C}{L} (\Delta Lk - Wr)^2 \quad (2.25)$$

donde C es la constante de torsión de la molécula y L es su longitud total [81]. Por lo tanto, para calcular la energía potencial elástica total de una molécula circular de DNA, con un determinado valor de ΔLk , es necesario conocer el valor de su Wr para cada conformación de prueba.

2.7.3 Selección de número de segmentos y rigidez de doblado

La elección del número de segmentos para la simulación de la molécula con el modelo WLC se relaciona con la cantidad de longitudes de Kuhn, n , contenidas a lo largo de la longitud de la molécula de DNA. La relación entre el número de segmentos N , y la cantidad de longitudes de Kuhn n contenidas en la molécula es $N = kn$ [80]. El número k se relaciona con la rigidez de doblado α mediante

$$k = \frac{1 + \langle \cos\theta \rangle}{1 - \langle \cos\theta \rangle} \quad (2.26)$$

donde

$$\langle \cos\theta \rangle = \frac{\int_0^\pi \cos\theta \sin\theta e^{-\alpha\theta^2} d\theta}{\int_0^\pi \sin\theta e^{-\alpha\theta^2} d\theta} \quad (2.27)$$

Las integrales son calculadas de forma numérica y el valor de α es utilizado para calcular la energía en la ecuación 2.23 [80][81].

2.8 Propiedades termodinámicas de los Intermediarios de Replicación

Un sistema se encuentra en estado de equilibrio termodinámico cuando es incapaz de cambiar de estado espontáneamente, o, sus variables macroscópicas como presión, temperatura y composición toman un valor constante, independiente del tiempo. El equilibrio es aquel estado al cual tiende espontáneamente un sistema [83]. Para ello ha de encontrarse simultáneamente en equilibrio térmico, mecánico y químico. Para que el sistema alcance equilibrio mecánico la fuerza neta en cada punto del sistema debe ser nula, del mismo modo es indispensable que el torque neto en todas partes del sistema sea cero. Como consecuencia de estas condiciones, un sistema alcanza equilibrio mecánico en un estado en el que no hay gradientes de energía entre las distintas partes del sistema.

Considerando a cada Intermediario de Replicación como un dominio topológico cerrado compuesto de dos partes, la región ya replicada y la no replicada, conectadas mecánicamente por las horquillas de replicación, y asumiendo que no interactúa con el medio circundante, entonces puede existir equilibrio termodinámico en el sistema. Para un Intermediario de Replicación se puede considerar a la energía potencial elástica como la energía interna del sistema, y el equilibrio termodinámico se da cuando los potenciales termodinámicos asociados al sistema, en particular la energía interna, se minimiza. De esta forma, cada exceso de estrés torsional en una región pasará a la otra en forma de energía elástica. Mecánicamente, un exceso de energía en la región no replicada, que se manifiesta en forma de un torque neto, inducirá la rotación de las horquillas de replicación, resultando en pérdida de cruces de superenrollamiento en la región no replicada, mientras va perdiendo estrés torsional, a la vez que la región ya replicada adquiere estrés torsional, el cual se manifiesta en forma de nuevos cruces de pre-encadenamiento (Figura 2.12A). De forma opuesta, un exceso de energía en la región ya replicada inducirá un torque neto en las horquillas de replicación, resultando en pérdida de cruces de pre-encadenamiento e introducción de nuevos cruces de superenrollamiento en la región no replicada. Este sistema estará en equilibrio

mecánico cuando el gradiente de energía entre ambas partes del RI sea igual a cero. Decimos entonces que el sistema está en equilibrio termodinámico cuando la energía total del RI es mínima, y ambas regiones del RI tienen la misma la energía elástica.

La región no replicada puede manifestar el efecto del estrés torsional mediante superenrollamiento (Figura 2.11 B - D), y la energía potencial elástica almacenada en esta región puede calcularse mediante las ecuaciones 2.23 y 2.25.

En condiciones *in vitro* la región ya replicada está permanentemente relajada debido a que los extremos libres de las cadenas recién sintetizadas pueden girar para liberar torsión. Por lo tanto, esta región no puede albergar estrés torsional en forma de superenrollamiento y la energía potencial elástica almacenada se puede calcular omitiendo los efectos de la energía de torsión. Es decir, la energía de esta región tiene solamente la componente de doblado E_b (ecuación 2.23).

Capítulo 3

Métodos para la simulación de moléculas circulares de DNA

3.1 Cálculo numérico de las propiedades topológicas

Las propiedades topológicas de las moléculas CCCs de DNA mencionadas hasta ahora, tales como el Lk y el Wr , pueden ser calculadas numéricamente para ser utilizadas en el modelo WLC. El Wr debe ser calculado para cada movimiento de prueba realizado por el método Metrópolis MC mientras que el concepto de Lk puede ser utilizado para calcular el número de enlaces entre dos moléculas, es decir, cuantas veces dos curvas se enrollan entre sí.

3.1.1 Cálculo del número de cruces de pre-encadenamiento

Para calcular la energía elástica de torsión del DNA basta con usar la ecuación 3.4. El modelo WLC requiere que el valor del ΔLk sea un parámetro constante en las simulaciones.

Sin embargo, es necesario tener un modo de cuantificar el grado de enrollamiento de una curva en torno a otra, como se da en el caso del pre-encadenamiento, esto sería el número de enlaces entre dos curvas. Para realizar el cálculo del número de enlaces se toma una proyección de las dos curvas enlazadas sobre un plano a elección. En particular, tomamos algún plano definido por los ejes principales de la estructura simulada (Figura 3.1).

La proyección sobre este plano permite localizar cruces de segmentos entre las dos curvas, así como la posición de cada segmento de la curva respecto al plano principal de proyección (para determinar los segmentos que están arriba y abajo en cada cruce).

El producto vectorial entre ambos segmentos dirigidos en un cruce permite discernir el tipo de cruce y por ende su signo. El número de cruces entre las cadenas nacientes de la región ya replicada se obtiene sumando los signos algebraicos de cada cruce y dividiendo el resultado entre dos.

Otro modo de calcular el número de enlace entre dos curvas está basado en la ecuación 2.4

$$n = \frac{1}{4\pi} \oint_{C'} \oint_C \left(\frac{(\mathbf{r}' - \mathbf{r}) \times d\mathbf{r}}{\|\mathbf{r}' - \mathbf{r}\|^3} \right) \cdot d\mathbf{r}', \quad (3.1)$$

De la ecuación 2.4 se puede inferir que si dos curvas están definidas por los vectores de posición, $\mathbf{r}$ y $\mathbf{r}'$ y la distancia entre dos segmentos consecutivos está dada por $\mathbf{b} = \mathbf{r}_{i+1} - \mathbf{r}_i$ y por $\mathbf{b}' = \mathbf{r}'_{j+1} - \mathbf{r}'_j$. Se elige tomar a estos dos vectores como los diferenciales de la ecuación 3.1, de tal manera que la aproximación se hace más exacta mientras menor sea el módulo de $\mathbf{b}$ y $\mathbf{b}'$ o, de forma relativa, si el número de segmentos en el modelo WLC se hace muy grande, entonces el número de enlaces de las dos curvas se puede calcular para N segmentos mediante:

$$n = \frac{1}{4\pi} \sum_{i=1}^N \sum_{j=1}^N \left(\frac{(\mathbf{r}'_j - \mathbf{r}_i) \times \mathbf{b}}{\|\mathbf{r}'_j - \mathbf{r}_i\|^3} \right) \cdot \mathbf{b}', \quad (3.2)$$

La implementación de un algoritmo que calcule la sumatoria doble de la ecuación 3.2 debe correr en tiempo polinomial $O(N^2)$.

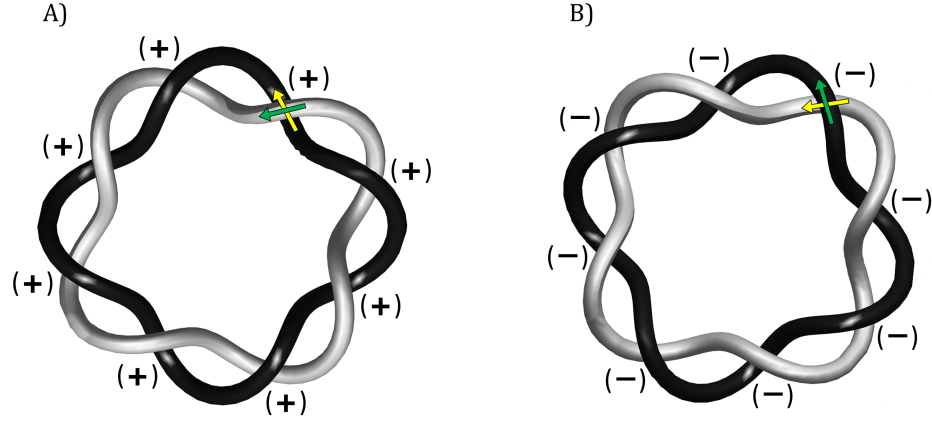


Figura 3.1: Determinación del número de enlaces mediante suma de signos topológicos. Se realiza una proyección bidimensional de los encadenados y se suman algebraicamente todos los valores asociados a los cruces encontrados (+1, -1). A) Encadenados con $Ca=+4$. B) Encadenados con $Ca=-4$.

3.1.2 Cálculo numérico del índice de superenrollamiento

A partir de la ecuación 2.19

$$Wr = \frac{1}{4\pi} \oint_C \oint_C \left(\frac{(\mathbf{r}_1 - \mathbf{r}_2) \times d\mathbf{r}_2}{\|\mathbf{r}_1 - \mathbf{r}_2\|^3} \right) \cdot d\mathbf{r}_1 \quad (3.3)$$

se pueden representar a los diferenciales mediante los vectores $\mathbf{b} = \mathbf{r}_{i+1} - \mathbf{r}_i$ y por $\mathbf{b}' = \mathbf{r}'_{j+1} - \mathbf{r}'_j$, y entonces podemos obtener la siguiente fórmula

$$Wr = \frac{1}{4\pi} \sum_{i=1}^N \sum_{j=1}^N \left(\frac{(\mathbf{r}_j - \mathbf{r}_i) \times \mathbf{b}_i}{\|\mathbf{r}_j - \mathbf{r}_i\|^3} \right) \cdot \mathbf{b}_j \quad (3.4)$$

A diferencia del cálculo del Lk , el cálculo del Wr requeriría que se recorra la misma curva dos veces, por ello se mide el ángulo sólido definido por cada par de segmentos dos veces. Para disminuir el tiempo de cómputo del Wr se puede recorrer la curva sólo una vez y luego se multiplica el resultado por dos. Esto se logra haciendo que la segunda suma comience desde $i + 1$, quedando la fórmula como:

$$Wr = \frac{1}{2\pi} \sum_{i=1}^N \sum_{j=i}^N \left(\frac{(\mathbf{r}_j - \mathbf{r}_i) \times \mathbf{b}_i}{\|\mathbf{r}_j - \mathbf{r}_i\|^3} \right) \cdot \mathbf{b}_j \quad (3.5)$$

Los algoritmos que calculen las sumatorias dobles de las ecuaciones 3.4 o 3.5 deben correr en tiempo polinomial $O(N^2)$.

3.1.3 Cálculo numérico del índice de torsión

Para el cálculo del índice de torsión recurrimos a la ecuación

$$Tw = \frac{1}{2\pi} \int_C \left(\frac{d\mathbf{a}}{ds} \times \hat{\mathbf{u}} \right) \cdot d\mathbf{u} \quad (3.6)$$

para las dos curvas de la región ya replicada se tienen en cuenta sus vectores de dirección, $\mathbf{r}$ y $\mathbf{r}'$, y el versor $\hat{\mathbf{u}}$ queda definido como $\hat{\mathbf{u}} = \frac{\mathbf{r} - \mathbf{r}'}{\|\mathbf{r} - \mathbf{r}'\|}$. La expresión numérica para la integral 3.6 es

$$Tw = \frac{1}{2\pi} \sum_{i=1}^N \frac{r_{i+1} - r_i}{\|r_{i+1} - r_i\|} \times \frac{r'_i - r_i}{\|r'_i - r_i\|} ((r'_{i+1} - r_{i+1}) - (r'_i - r_i)) \quad (3.7)$$

Un algoritmo que calcule la sumatoria de la ecuación 3.7 debe correr en tiempo polinomial $O(N)$.

3.2 Cálculo de la energía elástica

El cálculo de la energía de doblado de la ecuación 2.23 requiere la temperatura absoluta y la constante de doblado como parámetros iniciales. En estas simulaciones la temperatura de todo el sistema es la misma, con lo cual esta condición para alcanzar el equilibrio termodinámico se cumple. Obtener el valor de la constante de doblado, α , en el modelo WLC requiere el cálculo numérico de la ecuación 2.27 para cada combinación de valores de longitud de molécula y número de segmentos. Los valores de los ángulos de doblado entre dos segmentos consecutivos deben ser calculados en cada paso del procedimiento de simulación.

La energía potencial de torsión que se expresa en la ecuación 2.25 depende de la longitud de la molécula, la rigidez de torsión, y también de la variación del índice de ligamiento ΔLk como parámetros iniciales para la simulación. El Wr debe ser calculado en cada paso de la simulación, lo cual se realiza aplicando la ecuación 3.5 en cada paso del procedimiento de simulación.

Los términos de energía elástica de las ecuaciones 2.23 y 2.25 pueden ser utilizados para el cálculo de energía de cada conformación de prueba en el método Metrópolis MC. Los otros términos de energía, utilizados durante la prueba de energía que realiza el método Metrópolis MC, son el asociado al volumen excluido y la presencia de nudos. El valor de la energía aportado por estos términos suman cero o un valor lo suficientemente alto para evitar que la conformación pase la prueba de energía de necesario. Si dos segmentos están a una distancia mayor al diámetro efectivo y la conformación no presenta nudos, la energía que aportan es cero. En caso contrario, la energía que aportan tiene un valor lo suficientemente alto como para que la conformación sea rechazada. Para que esto ocurra, el valor debe ser mayor que la energía E_0 que tenía la conformación al inicio del procedimiento de simulación y además debe tener asociado un valor de $e^{\frac{(E_0 - E)}{k_b T}}$, menor a la probabilidad p de que la conformación sea aceptada por el método Metrópolis MC.

3.2.1 Término de energía asociado al diámetro efectivo

En la simulación con Metrópolis MC, con el modelo WLC, podemos asignar valores infinitos de energía potencial elástica a una conformación que tenga dos segmentos cualesquiera separados a una distancia menor o igual al diámetro efectivo ($d_{efectivo}$). Esta separación se calcula recorriendo la curva simulada dos veces, en el primer recorrido se toma un par de puntos consecutivos, con el segundo recorrido se toman otro par de puntos consecutivos. A continuación se halla la distancia entre los dos segmentos de las rectas definidas por los dos pares de puntos. Si en toda la curva no existe ningún par de segmentos a una distancia menor que $d_{efectivo}$, se suma cero a la energía total de la conformación. Si se encuentra al menos un par de rectas que tengan una distancia menor a $d_{efectivo}$ entre sí, se suma un valor lo suficientemente alto como para que la conformación sea rechazada.

Un algoritmo encargado de la verificación del $d_{efectivo}$ en una curva definida por un conjunto de vectores $r_1, \dots, r_i, \dots, r_N$ puede correr en tiempo polinomial $O(N^2)$ y la implementación de ese algoritmo para esta tesis se puede observar en el Anexo A.1.5.

3.3 Simulación numérica de Intermediarios de Replicación

En el presente trabajo se utilizó el modelo WLC para la simulación computacional de intermediarios de replicación con las horquillas de replicación detenidas una vez que la replicación ha avanzado hasta el 25, 50 y 75%. Los RIs de moléculas de DNA circular poseen una región no replicada y una región ya replicada, unidas entre sí por dos horquillas de replicación (Figura 3.2). En este modelo ambas regiones de la molécula están constituidas de curvas orientadas, representadas por los vectores de posición r_1 para la región no replicada, r_2 y r_3 para las dos cadenas nacientes de la región ya replicada. Estas tres curvas se unen en las dos horquillas de replicación. La longitud de la curva representada por r_1 respecto a la longitud de las curvas representadas por r_2 y r_3 , ambas del mismo tamaño, depende del porcentaje de avance de las horquillas de replicación en el RI simulado.

La aplicación del modelo WLC en esta obra utilizó dos tipos de movimientos para inducir cambios conformacionales en los RIs simulados. Uno de ellos es el de rotación rígida, el cual es un movimiento de prueba para el método Metrópolis MC, que puede ser aceptado o no debido a su efecto en el valor de la energía total. El otro movimiento es el de rotación de las horquillas de replicación, cuyo efecto se describe en la Figura 3.2. Además, se dotó a estos RIs de un valor constante de ΔLk como parámetro inicial.

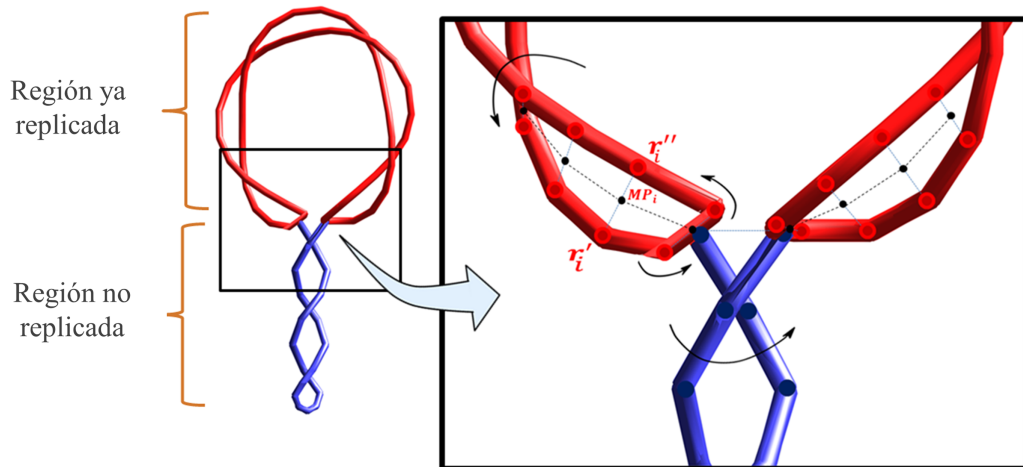


Figura 3.2: Intermediario de replicación obtenido con el modelo WLC. Modelo de molécula parcialmente replicada con superenrollamiento RH (-) en la región no replicada (azul) y pre-encadenamiento LH (-) en la región ya replicada (rojo). El mecanismo para modelar la rotación de las horquillas se muestra en la figura en recuadro. Puntos opuestos de las curvas replicadas, r'_i y r''_i , localizados en cada curva son rotados respecto al punto medio MP_i con el mismo ángulo φ . Las curvas ya replicadas se trenzan o destrenzan como resultado de la contribución de las rotaciones de todos los pares de puntos opuestos que conforman ambas curvas.

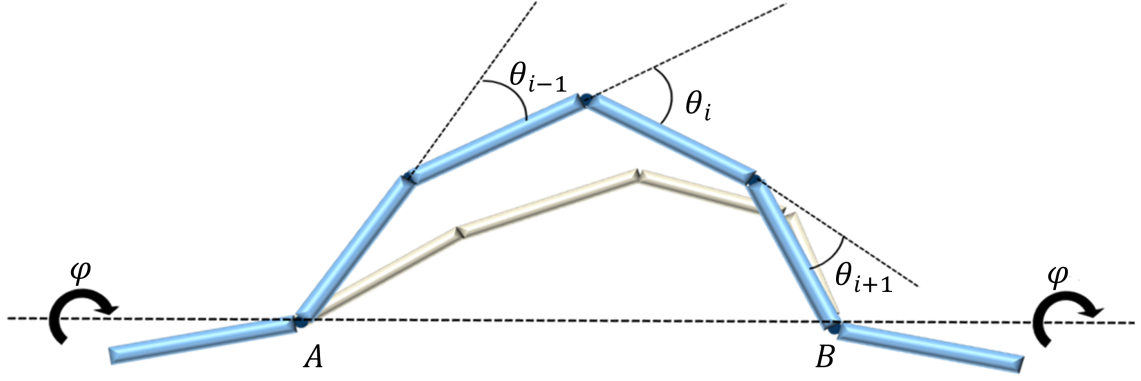


Figura 3.3: Movimiento de rotación rígida. El movimiento de rotación rígida o *crankshaft* consiste en la rotación de todos los vértices de los segmentos del modelo WLC en un mismo ángulo en torno a un eje que pasa por dos puntos aleatoriamente seleccionados (A y B). Este fue el movimiento de prueba utilizado para la aplicación del método Metrópolis MC en esta obra.

3.3.1 Movimiento de rotación rígida

El movimiento de rotación rígida o *crankshaft* se logra tomando aleatoriamente dos vértices de distintos segmentos, A y B , en la estructura simulada por el modelo WLC. Todos los puntos P contenidos entre A y B son rotados en torno a un eje que pasa por A y B en un ángulo φ . La elección de φ se realiza de manera aleatoria como parte del procedimiento de simulación de Metrópolis MC. El punto P_0 resultante de la rotación de P es

$$P_0 = [\hat{\mathbf{n}} \cdot (P - A)]\hat{\mathbf{n}} + \cos(\varphi)[(P - A) - [\hat{\mathbf{n}} \cdot (P - A)]\hat{\mathbf{n}}] \pm \sin(\varphi)[\hat{\mathbf{n}} \times (P - A)] \quad (3.8)$$

donde $\hat{\mathbf{n}}$ es el vector que apunta desde el punto A hasta el punto B , y $\hat{\mathbf{n}} = \frac{(B-A)}{\|B-A\|}$. Este movimiento de prueba se ilustra en la Figura 3.3. De manera aleatoria estos movimientos de rotación rígida pueden permitir que dos segmentos se crucen o atraviesen entre sí, provocando la aparición ocasional de nudos, causados por el procedimiento de simulación y no relacionados a fenómenos biológicos. Los nudos son detectados mediante el algoritmo basado en la búsqueda del polinomio de Alexander de la conformación [81].

3.3.2 Rotación de las horquillas de replicación

La topología de los RI se modifica enrollando o desenrollando la estructura plectonémica de la región no replicada y rotando las horquillas de replicación. El número de cruces de superenrollamiento en la región no replicada y el número de

cruces de pre-encadenamiento en la región ya replicada determinan la complejidad topológica del RI. El mecanismo para modelar la rotación de las horquillas de replicación se muestra en la Figura 3.2. La rotación de las horquillas se logra mediante el giro de las curvas correspondientes a las cromátidas hermanas una alrededor de la otra. En términos geométricos, esto implica la rotación de los dos puntos opuestos, $\mathbf{r}'_i$ y $\mathbf{r}''_i$ con respecto a su punto medio, $\mathbf{MP}_i$.

El segmento $\mathbf{r}'_i - \mathbf{MP}_i - \mathbf{r}''_i$ se gira sobre un ángulo φ desde su orientación original manteniendo el punto $\mathbf{MP}_i$ fijo. Dependiendo de la dirección de la rotación se obtiene como resultado que ambas curvas se trencen o destrencen, cambiando la topología de la región ya replicada. A través de este mecanismo, ambas horquillas pueden rotar en cada paso del procedimiento de simulación.

3.4 Método para la modificación de la topología de los Intermediarios de Replicación

Para las simulaciones de las propiedades energéticas y conformacionales de moléculas de DNA parcialmente replicadas mediante el modelo WLC se realizó el procedimiento que se describe a continuación. En primer lugar, se crearon las tres curvas representadas por los vectores de posición; r_1 para la región no replicada, r_2 y r_3 para la región ya replicada. El intermediario de replicación debe tener asociada una complejidad total caracterizada por ΔLk . En los casos en que todo el ΔLk de la molécula se encuentra localmente acumulado en la región no replicada, se debe dotar a esta región de superenrollamiento y esto se realiza aplicando en ella el método de Metrópolis MC [81]. Las curvas que definen a las nuevas cadenas complementarias de la región ya replicada también son sometidas al método Metrópolis MC. En este caso particular, un ejemplo de conformación inicial se muestra en la Figura 3.4. El número total de giros que realizará la horquilla de replicación, T , y el número total de pasos de la simulación, S , definen el ángulo de rotación que las horquillas efectúan en cada paso de la simulación $\sigma = 2\pi \frac{T}{S}$. La magnitud de la rotación total de las horquillas de replicación, $t = \frac{\sigma}{2\pi}$, determina la magnitud del cambio de complejidad de cada región para mantener constante el ΔLk en cada paso de la simulación.

Cada paso del procedimiento de simulación se divide en dos partes principales. En la primera parte la complejidad de cada región cambia en un valor absoluto t , o visto de otro modo, las horquillas giran en un ángulo σ . Si consideramos el caso en que inicialmente todo el estrés torsional se encuentra acumulado en la región no replicada y el número de paso en el que se encuentra la simulación es i , donde $i \in \mathbb{N}$ e $i =$

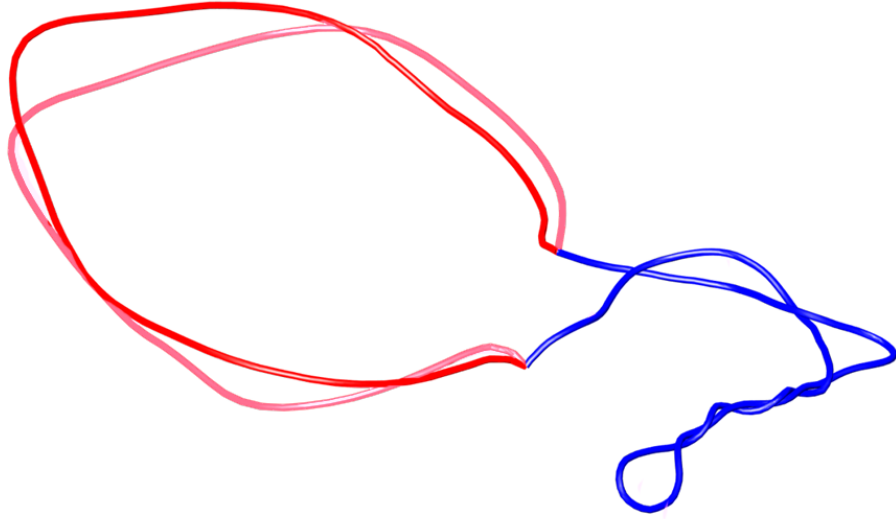


Figura 3.4: Conformación inicial un de Intermediario de Replicación simulado. Conformación correspondiente a una molécula que ha replicado el 50% del DNA. La región no replicada (azul) contiene todo el estrés torsional en forma de superenrollamiento LH (-). Las curvas correspondientes a las dos nuevas cadenas complementarias, de la región ya replicada (rojo) carecen de superenrollamiento por lo que están relajadas, sin pre-encadenamiento.

$1, 2, 3, \dots, S-1, S$. En general para cualquier paso, el valor de superenrollamiento en la región no replicada toma el valor $\Delta Lk - i \times t$, mientras que las horquillas de replicación ya habrán girado en un ángulo $i \times \sigma$ adquiriendo, o perdiendo, $i \times t$ vueltas. En el segunda parte de cada paso de la simulación se aplica el método Metrópolis MC a cada región del RI simulado para que cada región alcance la conformación más cercana posible al equilibrio, de modo que las regiones no replicada y ya replicada adquieren nuevas energías. Un ejemplo de cómo se aplican estos pasos de simulación se muestra en la Figura 3.5. La simulación continúa hasta alcanzar una complejidad seleccionada o una propiedad energética específica, (por lo que la simulación puede detenerse en algún caso en el que $i \neq S$). El algoritmo se configuró de modo que el aumento de pre-encadenamiento de las curvas replicadas cause una pérdida de superenrollamiento en la región no replicada, y viceversa.

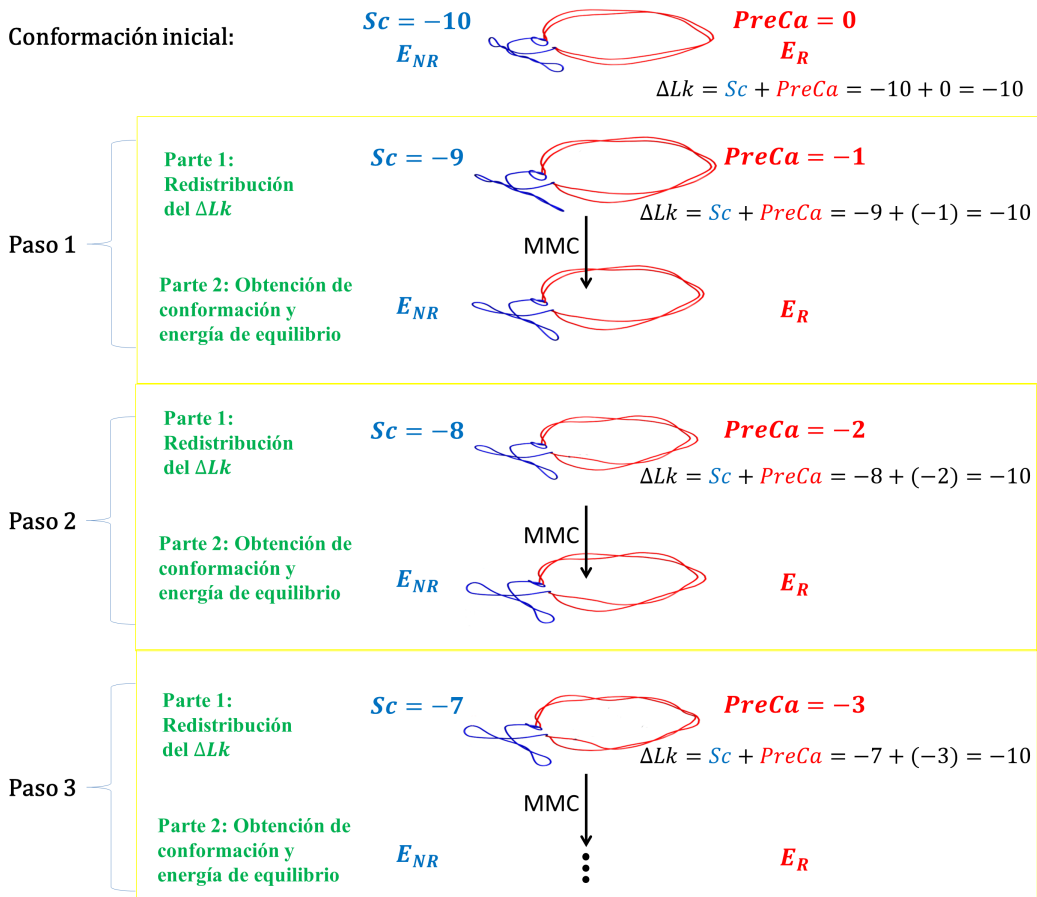


Figura 3.5: Procedimiento para modificar la topología de un Intermediario de Replicación. Procedimiento de la simulación para cambiar la conformación de un RI que alcanzó el 50% de la replicación. Cada paso de la simulación consta de 2 partes: En la primera, se cambia el valor del superenrollamiento de la región no replicada y se hacen girar las horquillas de replicación un número de vueltas igual a la variación en el superenrollamiento. En la segunda parte se utiliza el método Metrópolis Monte Carlo (MMC) para encontrar las nuevas conformaciones y energías E_{NR} y E_R de las regiones no replicada y ya replicada, respectivamente. Los pasos se repiten hasta encontrar una conformación específica o una propiedad energética particular.

3.5 Procedimiento de simulación numérica

Utilizando el modelo WLC se simularon RIs con horquillas detenidas que replicaron el 25, 50 ó 75% de la replicación del DNA, dependiendo de la distancia las horquillas con respecto al origen de replicación.

En este modelo el DNA se representa como una serie de segmentos cilíndricos rígidos de longitud constante de 3.4 nm y con diámetro específico de 3.0 nm, impenetrables entre sí.

Las moléculas de DNA circular simuladas tenían un tamaño de 4440 pb, equivalente a 400 segmentos. Para el caso de moléculas que alcanzaron el 50% de la replicación, la región no replicada estaba compuesta de 200 segmentos (2220 pb) y cada una de las nuevas cadenas complementarias de la región ya replicada constaba de 200 segmentos.

En el caso de moléculas replicadas hasta el 25%, la región no replicada constaba de 300 y las nuevas cadenas complementarias de la región ya replicada estaba formada por 100 segmentos, mientras que en las moléculas replicadas hasta el 75%, la región no replicada tenía 100 segmentos y la ya replicada 300 segmentos.

Se dotó a todas las conformaciones simuladas de un ΔLk que permaneció constante durante todo el proceso la simulación. En las conformaciones iniciales el 100% o el 80% del ΔLk se acumuló en una sola región del RI mientras que la otra región tenía una conformación más relajada.

Para estudiar el efecto que tiene la conformación en las propiedades termodinámicas se realizaron simulaciones en las que todo el estrés torsional, que estaba concentrado en una región del RI, pasó completamente a la otra región. Este tipo de simulación se realizó sólo para moléculas que replicaron el 50% del DNA, de dos maneras distintas;

- 10 simulaciones independientes de RIs con todo el ΔLk inicialmente concentrado en la región no replicada, en forma de superenrollamiento, y la región ya replicada completamente relajada, sin pre-encadenados.
- 10 simulaciones independientes de RIs con todo el ΔLk inicialmente concentrado en la región ya replicada, en forma de pre-encadenamiento, mientras que la región no replicada se encontraba relajada, sin superenrollamiento.

Por otro lado, también se analizaron las propiedades conformacionales de un RI en el estado de equilibrio termodinámico. Para ello se realizaron simulaciones en las que el estrés torsional pasó de la región con mayor ΔLk a la región con menor estrés torsional, deteniendo la simulación cuando el gradiente de energía entre ambas se anuló, obteniéndose así la distribución de equilibrio del ΔLk . Este estudio se realizó de la siguiente forma;

- 10 simulaciones independientes de moléculas que replicaron la mitad del DNA, en las que todo el ΔLk inicialmente se encontraba en la región no replicada en forma de superenrollamiento y la región ya replicada libre de pre-encadenamiento. El estrés torsional de la región no replicada migró a la ya replicada hasta alcanzar el equilibrio termodinámico. El procedimiento de simulación se realizó a partir de cuatro conformaciones iniciales: 1) RIs con superenrollamiento negativo, $\Delta Lk = -10$. 2) RIs con superenrollamiento

negativo, $\Delta Lk = -20$. 3) RIs con superenrollamiento positivo, $\Delta Lk = +10$. 4) RIs con superenrollamiento positivo, $\Delta Lk = +20$.

- 10 simulaciones independientes de moléculas en las que la replicación avanzó hasta el 50%, con el ΔLk distribuido asimétricamente entre ambas regiones. El procedimiento de simulación se realizó a partir de dos conformaciones iniciales: 1) RIs con 8 cruces de superenrollamiento negativo en la región no replicada y 4 cruces de pre-encadenamiento negativo, $\Delta Lk = -10$. 2) RIs con 2 cruces de superenrollamiento negativo en la región no replicada y 16 cruces de pre-encadenamiento negativo, $\Delta Lk = -10$.
- 10 simulaciones de moléculas donde la replicación avanzó hasta el 25%, con todo el ΔLk inicialmente acumulado en la región no replicada en forma de superenrollamiento negativo y la región ya replicada libre de pre-encadenamiento, $\Delta Lk = -10$
- 10 simulaciones de moléculas donde la replicación avanzó hasta el 75%, con todo el ΔLk inicialmente acumulado en la región no replicada en forma de superenrollamiento negativo y la región ya replicada sin pre-encadenamiento, $\Delta Lk = -10$

El modelado de los RIs, la implementación de la rotación de las horquillas de replicación y el método de Metrópolis MC se realizaron mediante el programa MATLAB R2016b, en un ordenador con procesador Inter(R) core(TM) i5-8250U de 1.60 - 1.80 GHz, con memoria de 4GB. En la Tabla 3.1 mostramos las constantes utilizadas.

Tabla 3.1: Constantes utilizadas en la simulaciones realizadas

Símbolo	Nombre	Valor
k_B	Constante de Boltzmann	$1.38^{-23} \frac{J}{K}$
T	Temperatura absoluta	$297 K$
C	Constante de torsión	$4.50^{-19} j.pb$
α	Constante de doblado	6.53
$2a$	Longitud de Kuhn	$295 pb$

Capítulo 4

Resultados

La presente obra se centra en el estudio de la coordinación entre superenrollamiento y pre-encadenamiento durante la replicación del DNA y sus propiedades termodinámicas. Las simulaciones se basaron en moléculas de DNA circular con un tamaño de 4440 pb con las horquillas de replicación detenidas. El presente trabajo es un estudio mediante simulación computacional de la distribución de estrés torsional en moléculas de DNA circular parcialmente replicadas con las horquillas de replicación detenidas en distintos puntos con relación al origen de replicación. Con el fin de llevar a cabo este análisis se utilizó el modelo *worm-like chain* para crear Intermediarios de Replicación que constan de una región no replicada y otra ya replicada. Al modelo obtenido se le aplicó un sistema de rotación de las horquillas de replicación para representar el fenómeno de la migración del estrés torsional de una región a otra del RI. Para las simulaciones numéricas se generó un conjunto de moléculas parcialmente replicadas en las que la replicación ha avanzado hasta el 25, 50 ó 75% de la molécula de DNA. Las conformaciones iniciales fueron sometidas a un estrés torsional local generando un desbalance en el torque. Las conformaciones de equilibrio termodinámico se encontraron mediante el protocolo de Metrópolis MC.

4.1 Propiedades termodinámicas en función de las características conformacionales de los Intermediarios de Replicación

Para estudiar el intercambio de estrés torsional entre las regiones no replicada y ya replicada de moléculas de DNA parcialmente replicadas en condiciones *in vitro*, realizamos simulaciones numéricas utilizando el modelo WLC y el método Metrópolis

MC para analizar la distribución de energía potencial elástica en RIs 50% replicados migrando todo el estrés torsional de una región a otra.

En primer lugar, modelamos RIs con 10 cruces de superenrollamiento RH (-) y sin cruces de pre-encadenamiento. Todo el estrés torsional se acumuló en la región no replicada en forma de superenrollamiento, mientras que en la región ya replicada las cadenas hermanas estaban relajadas. Este tipo de conformación se observa en la Figura 3.4.

En nuestras simulaciones las horquillas de replicación se rotaron para permitir la migración de todos los cruces RH (-) de la región no replicada a la región ya replicada. La Figura 4.1 muestra la variación de energía después de cada rotación de las horquillas en función al número de cruces en las regiones ya replicada y no replicada. Se observa claramente que el valor de la energía elástica debida al superenrollamiento del DNA en la región no replicada disminuye a medida que aumenta el número de pre-encadenados en la región ya replicada. La energía potencial elástica promedio, calculada a partir de 10 simulaciones independientes, de la conformación inicial fue de 34.2 Kcal/mol, 29.6 Kcal/mol para la región no replicada y 4.7 Kcal/mol para la porción ya replicada. Al final, la energía elástica de la región no replicada se redujo a 4.2 Kcal/mol mientras que la energía elástica de la región replicada aumentó a 46.6 Kcal/mol. Ambas regiones de la molécula tuvieron la misma energía, de 12.6 Kcal/mol cada una, para una conformación en que la región no replicada tenía 4.7 cruces de superenrollamiento RH (-) y la región ya replicada 5.3 pre-encadenados LH (-).

También se modeló la situación opuesta, es decir, RIs 50% replicados con 20 cruces de pre-encadenamiento LH (-) en la región ya replicada y sin superenrollamiento en la región no replicada. En este caso, el giro de las horquillas de replicación permitió la migración de todo el estrés torsional desde la región ya replicada a la región no replicada. Los valores iniciales de energía elástica, resultantes de 10 simulaciones independientes, fueron 2.9 y 50.2 Kcal/mol para las regiones no replicada y ya replicada, respectivamente. Las horquillas se giraron para permitir la migración de todos los cruces LH (-) de la región ya replicada a la no replicada. La Figura 4.2 muestra la variación de energía después de cada rotación en función al número de cruces en ambas regiones. Se observa claramente que el valor de la energía potencial elástica debida al pre-encadenamiento en la región ya replicada disminuyó a medida que aumentó el superenrollamiento en la región no replicada. Al final, la energía potencial elástica de la región replicada se redujo a 12,6 Kcal/mol, mientras que la energía elástica de la región replicada aumentó a 35,6 Kcal/mol. El punto en que ambas regiones de la molécula tuvieran la misma energía, cada una de ellas con 18.3 Kcal/mol, correspondió a una conformación en que la región no replicada tiene 4.5 cruces de superenrollamiento RH (-) y la región ya replicada tiene 5.5 pre-encadenados LH (-).

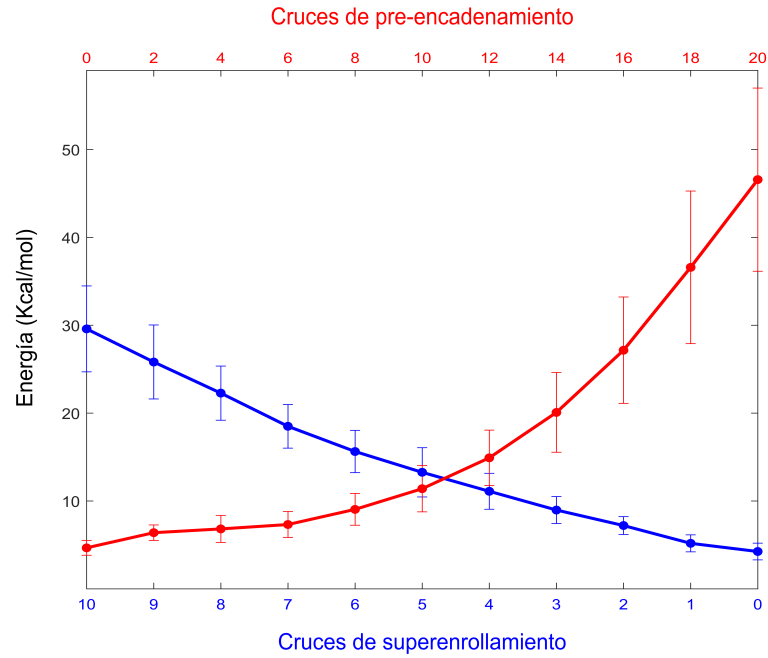


Figura 4.1: Migración del estrés torsional de la región no replicada a la ya replicada de un RI que alcanzó el 50% de la replicación. Las horquillas de replicación rotaron hasta pasar todo el estrés torsional desde la región no replicada a la ya replicada ($\Delta Lk = -10$). Energía elástica de la región no replicada (línea azul) y la región ya replicada (línea roja) en función al número de cruces de cada región. Valores promedio de la energía elástica de 10 simulaciones independientes.

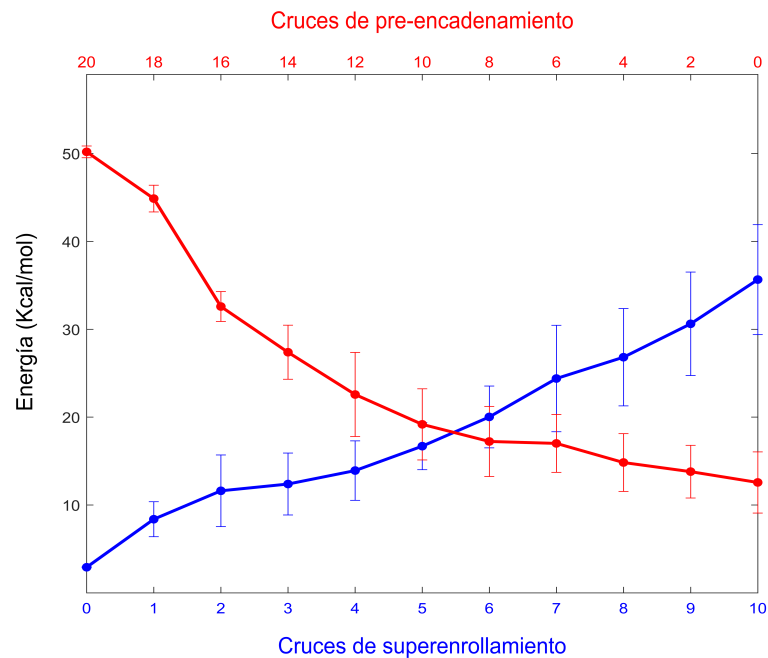


Figura 4.2: Migración del estrés torsional de la región ya replicada a la no replicada de un RI que alcanzó el 50% de la replicación. Las horquillas de replicación rotaron hasta pasar todo el estrés torsional ($\Delta Lk = -10$) desde la región ya replicada a la no replicada. Variación de la energía elástica de la región no replicada (línea azul) y la región ya replicada (línea roja) en función al número de cruces de superenrollamiento y pre-encadenamiento. Valores promedio de 10 simulaciones independientes.

Estos resultados demostraron que nuestro modelo para un RI en condiciones *in vitro* es capaz de reproducir la migración del estrés torsional desde la región no replicada hacia la región ya replicada y viceversa. El grado de acumulación del estrés torsional en cada región fue controlado de tal modo que cada sistema simulado pasó por un proceso no espontáneo. Queda evidenciada mediante estos resultados la existencia de una relación entre la energía elástica y la deformación torsional para ambas regiones del RI.

4.2 Características conformacionales en función de las propiedades termodinámicas de los Intermediarios de Replicación

En este apartado del estudio buscamos conocer la conformación que presenta un RI una vez que este se encuentra en el estado de equilibrio termodinámico. Las conformaciones iniciales fueron sometidas a un estrés torsional local generando un desbalance en el torque. El estrés torsional se distribuyó entre ambas regiones, mediante la rotación de las horquillas de replicación, hasta que el gradiente de energía entre ambas se anuló, encontrándose de este modo la conformación de equilibrio. Esto se realizó para moléculas de DNA circular con un avance del 25, 50 ó 75% de las horquillas de replicación.

4.2.1 Simulación de Intermediarios de Replicación con un avance del 25% de las horquillas de replicación

Con el fin de estudiar condiciones análogas a las de experimentos realizados por Cebrian et al. [1], analizamos la distribución de la energía elástica en RIs en los cuales las horquillas de replicación están detenidas y la región ya replicada corresponde al 25% del tamaño de la molécula. Se realizaron 10 simulaciones independientes en las cuales se dotó a la región no replicada de 10 cruces de superenrollamiento LH (-), mientras que la región ya replicada se encontraba relajada, sin cruces de pre-encadenamiento. La rotación de las horquillas de replicación permitió la migración del estrés torsional hasta llegar a un estado en que el torque neto en las horquillas de replicación es nulo. La variación de energía con cada rotación de las horquillas se muestra en la Figura 4.3.

Para el caso promedio, resultante de 10 simulaciones independientes, en la conformación inicial la energía elástica de la región ya replicada fue de 6.1 Kcal/mol y se incrementó hasta 12.5 Kcal/mol al llegar a la conformación de equilibrio, mientras que en la región no replicada fue decreciendo de su valor inicial de 21.1 Kcal/mol hasta 12.5 Kcal/mol. La conformación de equilibrio en la que el torque neto desaparece correspondió a un RI con

6.8 cruces de superenrollamiento RH (-) en la región no replicada y 3.2 pre-encadenados LH (-) en la región ya replicada.

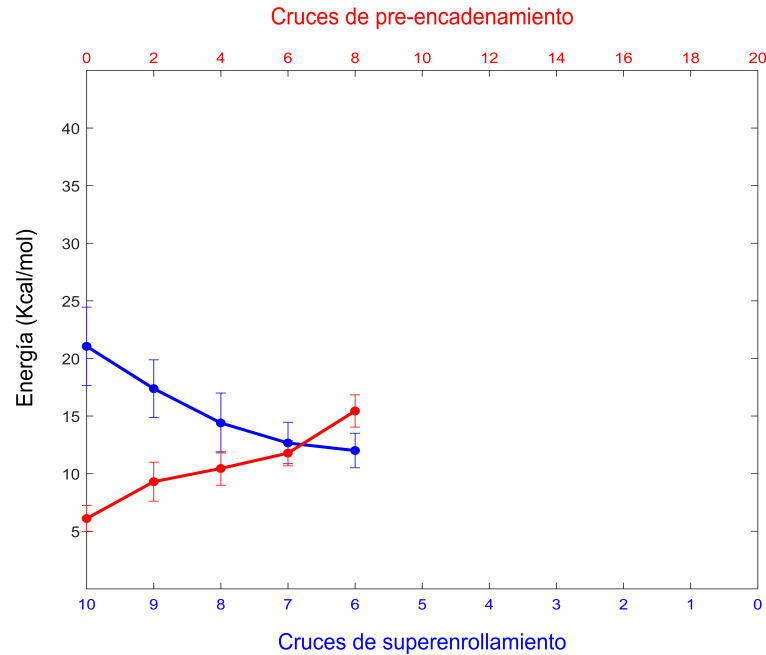


Figura 4.3: Distribución del estrés torsional en RIs con un avance del 25% de las horquillas de replicación. Se partió de una conformación inicial con 10 cruces de superenrollamiento LH (-), mientras que la región ya replicada se encontraba relajada, ($\Delta Lk = -10$). Las horquillas de replicación rotaron hasta que el RI alcanzó el estado de equilibrio termodinámico.

Los resultados de estas simulaciones demostraron que, en el estado de equilibrio termodinámico, el estrés torsional de moléculas que han replicado un 25% de su DNA se distribuye de tal modo que la mayor parte del ΔLk queda albergado en la región no replicada, en forma de superenrollamiento.

4.2.2 Simulación de Intermediarios de Replicación con un avance del 50% de las horquillas de replicación

4.2.2.1 Intermediarios de Replicación con superenrollamiento negativo

Creamos un conjunto de conformaciones de RIs de 4440 pb que han replicado la mitad del DNA. Primeramente, se simularon RIs con todo el estrés torsional acumulado en la región no replicada. Esta conformación es topológicamente idéntica a la representada en la Figura 3.4. El estrés torsional se encontraba acumulado en la región no replicada en forma de 10 cruces de superenrollamiento RH (-) mientras que la región ya replicada estaba relajada, sin pre-encadenamiento. La rotación de las horquillas de replicación

permitió el paso del estrés torsional hacia la región ya replicada hasta que la molécula alcanzó el equilibrio termodinámico, es decir, la conformación en la cual el torque neto de la molécula desaparece y el RI alcanza el equilibrio mecánico.

La variación energía elástica en función al número de cruces de ambas regiones se muestra en la Figura 4.4 A. La conformación inicial, promedio de 10 simulaciones independientes, presentó energías iniciales de 27.2 y 4.7 Kcal/mol para las regiones no replicada y ya replicada respectivamente. En el estado de equilibrio termodinámico, el valor final de la energía elástica fue de 13.7 Kcal/mol para ambas regiones. Bajo estas condiciones, se encontró que las moléculas que han replicado el 50% del DNA alcanzaron el equilibrio termodinámico migrando la mitad del superenrollamiento para formar la misma cantidad de pre-encadenados. La conformación de equilibrio correspondió a un RI con 5.3 cruces de superenrollamiento RH (-) en la región no replicada y 4.7 pre-encadenados LH (-).

Con el objeto de analizar el efecto de la complejidad topológica, se crearon RIs que han replicado el 50% su DNA, con todo el estrés torsional inicialmente confinado en la región no replicada. Se simularon RIs con un $\Delta Lk = -20$, cuya conformación inicial correspondía a un RI con 20 cruces de superenrollamiento LH (-) y ausencia de pre-encadenados en la región ya replicada. Tal como en el caso previo, la rotación de las horquillas de replicación permitió la migración parcial del estrés torsional almacenado en la región no replicada hasta que la molécula alcanzó el equilibrio termodinámico. La energía elástica de ambas regiones en función al número de cruces se muestra en la Figura 4.4 B. Para el caso promedio de 10 simulaciones independientes, la conformación inicial tenía valores de energía de 135.2 y 8.0 Kcal/mol para las regiones no replicada y ya replicada, respectivamente. En el estado de equilibrio termodinámico el valor final de la energía elástica fue de 40.7 Kcal/mol para cada región. Bajo estas condiciones, las moléculas que han replicado el 50% del DNA alcanzaron el equilibrio termodinámico convirtiendo aproximadamente la mitad de los cruces de superenrollamiento de la región no replicada en pre-encadenados. La conformación con torque nulo en las horquillas de replicación correspondió a un RI con 11.2 cruces de superenrollamiento RH (-) en la región no replicada y 8.8 pre-encadenados LH (-) en la región ya replicada.

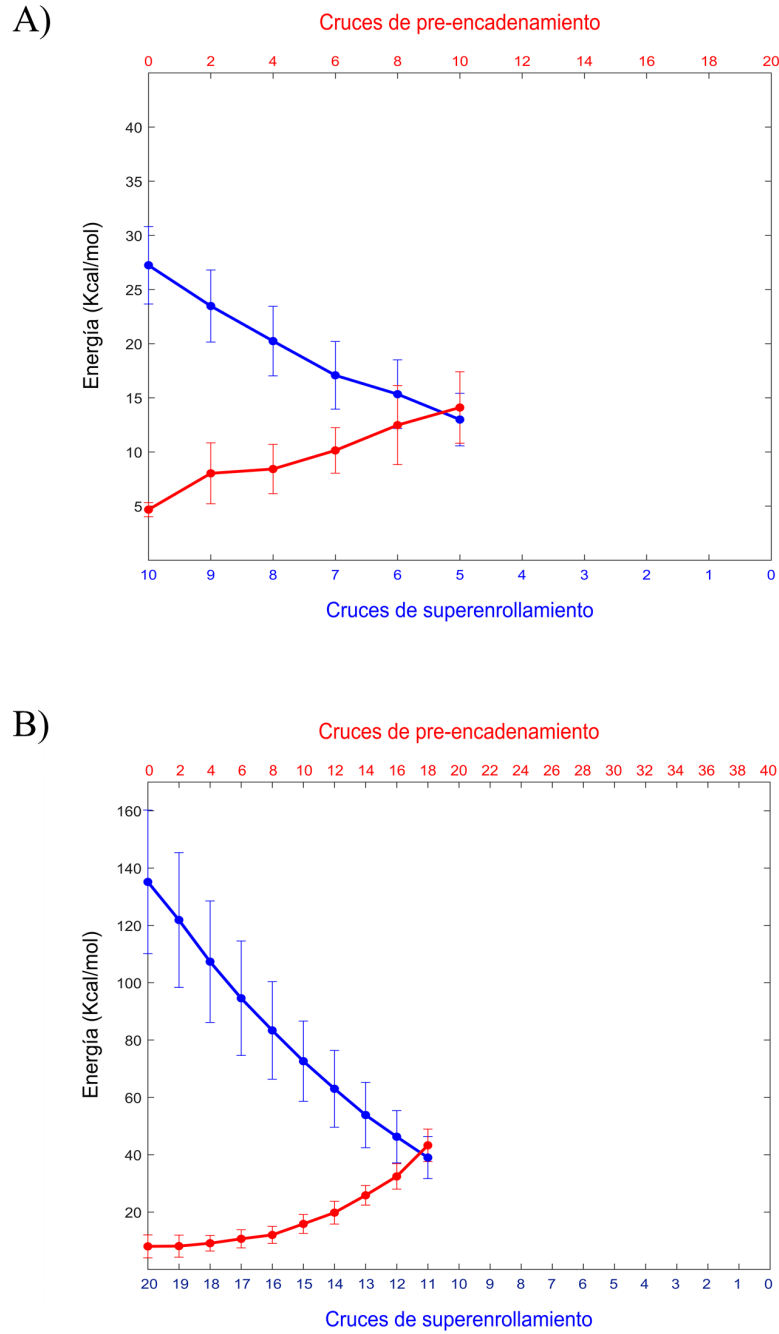


Figura 4.4: Distribución del estrés torsional en RIs que han replicado un 50% de su DNA y acumulan superenrollamiento negativo. Las horquillas de replicación rotaron hasta que el RI alcanzó el estado de equilibrio termodinámico. A) Variación de la energía elástica en RIs 50% replicados con $\Delta Lk = -10$. B) Variación de la energía elástica en RIs 50% replicados con $\Delta Lk = -20$. Se muestra el promedio de la energía elástica para las regiones no replicada (línea azul) y ya replicada (línea roja) correspondiente a 10 simulaciones independientes.

Para extender estos estudios a casos en los que el estrés torsional de la conformación inicial se encuentra distribuido de manera asimétrica entre ambas regiones del RI, se simularon RIs con 8 cruces de superenrollamiento RH (-) en la región no replicada y

4 cruces de pre-encadenamiento LH (-) en la región ya replicada, $\Delta Lk = -10$. En el procedimiento de simulación la rotación de las horquillas permitió la migración del estrés torsional de la región con mayor energía elástica acumulada hacia la región con menor energía, de modo que el torque en la molécula se equilibre. La variación de energía elástica en función al número de cruces en ambas regiones se presenta en la Figura 4.5 A. Se observa la pérdida gradual de cruces en la región no replicada, con el consecuente aumento de pre-encadenados entre las cadenas de DNA ya replicadas. Los valores iniciales de energía elástica para las regiones no replicada y ya replicada fueron 21.9 y 7.7 Kcal/mol, respectivamente. La rotación de las horquillas paró cuando ambas regiones alcanzaron la misma energía elástica. La conformación de equilibrio, resultante de 10 simulaciones independientes, correspondió a un RI con 5.1 cruces de superenrollamiento RH (-) en la región no replicada y 4.9 pre-encadenados LH (-) en la región ya replicada.

Por otro lado, con el fin de estudiar la conformación inicial opuesta, se modeló un conjunto de RIs con 2 cruces de superenrollamiento RH (-) en la región no replicada y 16 cruces de pre-encadenamiento LH (-), equivalentes a 8 pre-encadenados (-), en la región ya replicada, $\Delta Lk = -10$. El procedimiento de simulación permitió la rotación de las horquillas de replicación de modo que el estrés torsional migró hacia la región más relajada. El procedimiento se detuvo cuando los RIs simulados alcanzaron el estado de equilibrio mecánico. La conformación de equilibrio termodinámico, promedio de 10 simulaciones independientes, correspondió a un RI con 5.3 cruces de de superenrollamiento RH (-) y 4.7 pre-encadenados LH (-) (Figura 4.5 B). Los valores iniciales de energía para las regiones no replicada y ya replicada fueron 6.7 y 24.1 Kcal/mol, respectivamente. La energía elástica total de la conformación de equilibrio fue de 27.2 Kcal/mol, 13.6 Kcal/mol para cada región.

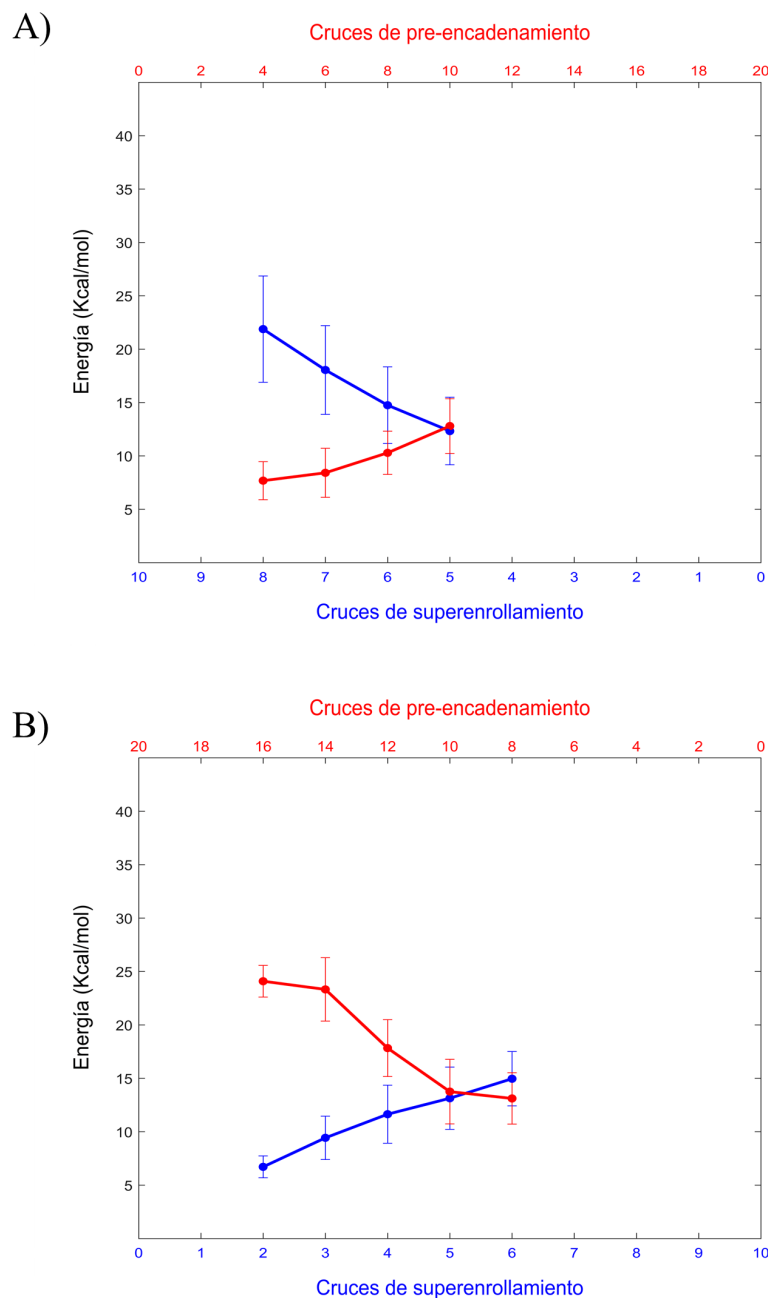


Figura 4.5: RIs que han replicado un 50% de su DNA y presentan una distribución asimétrica del estrés torsional. Las horquillas de replicación rotaron hasta que ambas regiones alcanzaron el equilibrio termodinámico. A) Variación de energía elástica en RIs con 8 cruces de superenrollamiento RH (-) y 2 cruces de pre-encadenamiento LH (-), $\Delta Lk = -10$. B) Variación de la energía elástica en RIs con 2 cruces de superenrollamiento RH (-) y 8 cruces de pre-encadenamiento LH (-). Se muestra el promedio de la energía elástica de 10 simulaciones independientes para la región no replicada (línea azul) y la ya replicada (línea roja).

En conjunto, estos resultados demuestran que para todas las conformaciones iniciales de moléculas con 50% del DNA replicado la rotación las horquillas de replicación permitió alcanzar el estado de equilibrio termodinámico. En todos los casos, independientemente

de la conformación inicial, en el estado de equilibrio se obtuvo una distribución simétrica del estrés torsional entre ambas regiones.

4.2.2.2 Intermediarios de Replicación con superenrollamiento positivo

Como se mencionó previamente, durante la replicación del DNA *in vivo*, el desenrollamiento de la doble hélice parental de DNA, causado por la DNA helicasa, produce un estrés torsional compensatorio por delante de las horquillas de replicación que se manifiesta en forma de superenrollamiento LH (+). La rotación de las horquillas de replicación reduce el estrés torsional (+) acumulado en la región no replicada, generando pre-encadenamientos (+) en la región ya replicada.

Con el fin de ampliar nuestro análisis a las condiciones *in vivo*, se modelaron RIs con estrés torsional (+). En la conformación inicial se dotó a la región no replicada de 10 cruces de superenrollamiento LH (+) mientras que la región ya replicada carecía de pre-encadenamiento. Mediante la rotación de las horquillas de replicación se permitió la migración parcial del estrés torsional acumulado en la región no replicada hasta alcanzar el estado de equilibrio termodinámico. La energía elástica de ambas regiones en función a la complejidad topológica se muestra en la Figura 4.6 A. El promedio de 10 simulaciones independientes correspondió a una conformación inicial con valores de energía de 25.6 y 5.7 Kcal/mol para las regiones no replicada y ya replicada, respectivamente. En el estado de equilibrio termodinámico el valor final de la energía fue de 11.4 Kcal/mol en cada región. Bajo estas condiciones, los RIs alcanzaron el equilibrio termodinámico migrando, a través de las horquillas de replicación, la mitad de los cruces de la región no replicada para formar el mismo número de cruces de pre-encadenamiento RH (+). La conformación promedio de equilibrio corresponde al de un RI con 5.4 cruces de superenrollamiento LH (+) en la región no replicada y 4.6 pre-encadenados RH (+).

Repetimos el procedimiento de simulación para RIs con la mitad del DNA replicado, a los cuales se dotó de 20 cruces de superenrollamiento LH (+) en la región no replicada y ningún pre-encadenamiento. Como en el caso previo, el giro de las horquillas de replicación permitió la migración parcial del estrés torsional para alcanzar el equilibrio termodinámico. La variación de energía elástica de ambas regiones en función al número de cruces en ambas regiones se muestra en la Figura 4.6 B. El valor promedio de la energía elástica inicial de 10 simulaciones independientes fue de 120.5 y 6.2 Kcal/mol para las regiones no replicada y replicada, respectivamente. En el estado de equilibrio termodinámico el valor de la energía elástica fue de 35.8 Kcal/mol para cada región. Bajo estas condiciones, las moléculas parcialmente replicadas alcanzan el equilibrio termodinámico migrando, a través de las horquillas de replicación,

aproximadamente la mitad de los cruces de superenrollamiento LH (+) para formar un número similar de cruces de pre-encadenamiento RH (+). La conformación de equilibrio, resultante de 10 simulaciones independientes, corresponde a un RI con 10.8 cruces plectonémicos LH (+) y 9.2 pre-encadenados RH (+).

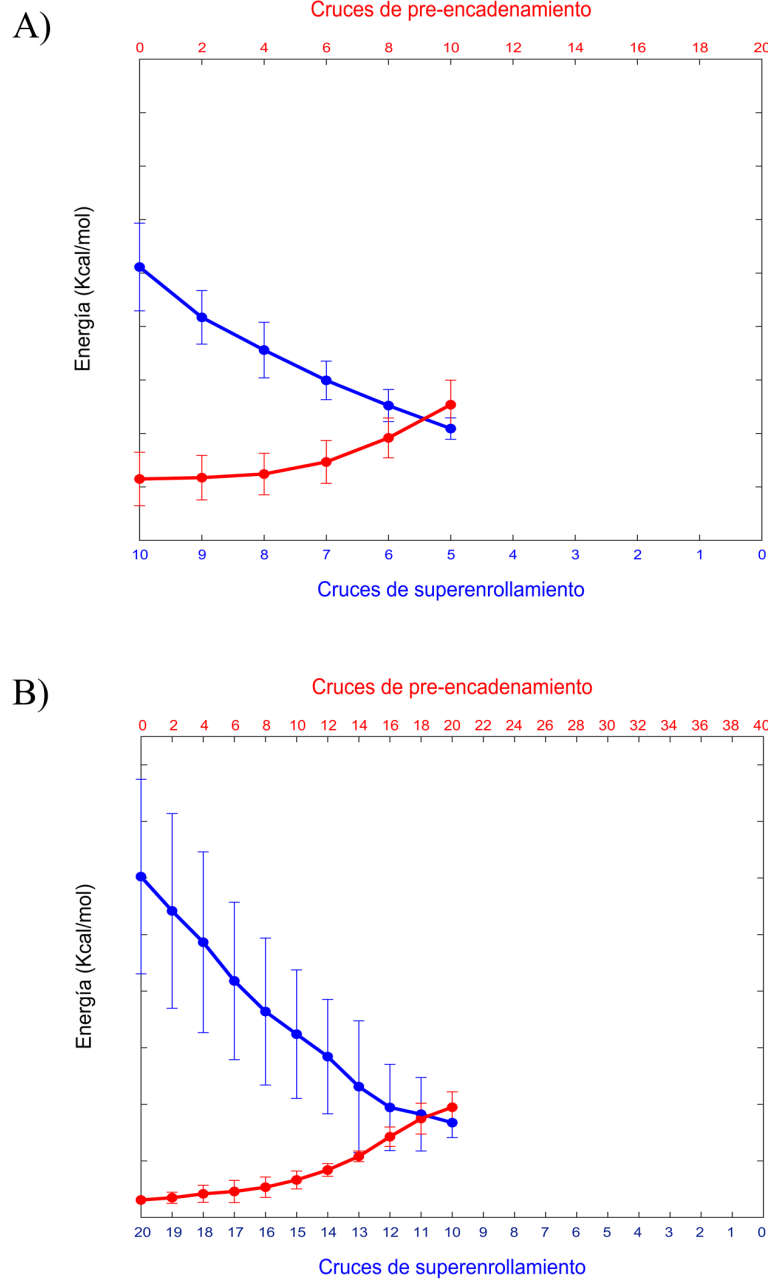


Figura 4.6: Distribución del estrés torsional en RIs que han replicado un 50% de su DNA y acumulan superenrollamiento positivo. Las horquillas de replicación rotaron hasta que el RI alcanzó el estado de equilibrio termodinámico. A) Variación de la energía elástica en RIs 50% replicados con $\Delta Lk = +10$. B) Variación de la energía elástica de RIs con $\Delta Lk = +20$. Se muestra el promedio de la energía elástica de 10 simulaciones independientes para la región no replicada (línea azul) y la ya replicada (línea roja).

Estos resultado demuestran que los RIs replicados hasta el 50% con estrés torsional acumulado en forma de superenrollamiento positivo alcanzan el estado de equilibrio termodinámico cuando cada región de la molécula queda con aproximadamente la mitad del ΔLk , de manera similar a lo observado para los RIs replicados hasta el 50% con estrés torsional acumulado en forma de superenrollamiento negativo.

4.2.3 Simulación de Intermediarios de Replicación con un avance del 75% de las horquillas de replicación

Tal como se mencionó en la sección 4.2.1, con el fin de analizar el efecto del avance de las horquillas de replicación en la distribución del estrés torsional, se modelaron RIs con las horquillas detenidas en distintos puntos antes de la terminación de la replicación. En este caso, se modelaron RIs con las horquillas detenidas que han replicado el 75% del DNA. Se realizaron 10 simulaciones independientes en las cuales se dotó a la región no replicada de 10 cruces de superenrollamiento RH (-), mientras que la región ya replicada carecía de pre-encadenamiento. La rotación de las horquillas de replicación para alcanzar el equilibrio termodinámico provocó la migración del superenrollamiento RH (-) hacia la región ya replicada, formando pre-encadenados LH (-). La variación de la energía elástica en ambas regiones, de un promedio de 10 simulaciones independientes, se muestra en la Figura 4.7. La rotación de las horquillas se detuvo cuando ambas regiones alcanzaron el equilibrio termodinámico, en el cual la energía elástica de cada región fue de 15.6 Kcal/mol. La conformación de equilibrio corresponde a un RI con 3.5 cruces de superenrollamiento RH (-) en la región no replicada y 6.5 cruces de pre-encadenamiento LH (-) en la región ya replicada.

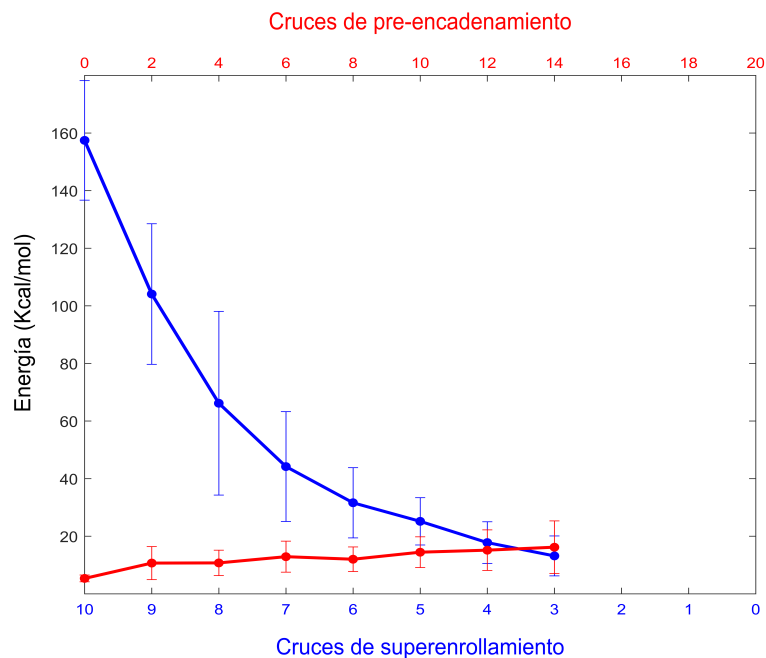


Figura 4.7: Distribución del estrés torsional en RIs con un avance del 75% de las horquillas de replicación. Las horquillas de replicación rotaron hasta que el RI alcanzó el estado de equilibrio termodinámico. Variación de la energía elástica de la región no replicada (línea azul) y la región ya replicada (línea roja) en función al número de cruces de superenrollamiento y pre-encadenamiento. Valores promedio de 10 simulaciones independientes.

Estos resultados demuestran que, para las moléculas en las cuales la horquilla de replicación ha avanzado hasta el 75%, en el estado de equilibrio termodinámico, la mayor parte del estrés torsional queda almacenado en la región ya replicada.

4.2.4 Distribución del estrés torsional en función al tamaño relativo de ambas regiones del Intermediario de Replicación

En la Figura 4.8 se muestran las conformaciones iniciales y finales de RIs con un avance del 25, 50 y 75% de las horquillas de replicación, todos con $\Delta Lk = -10$.

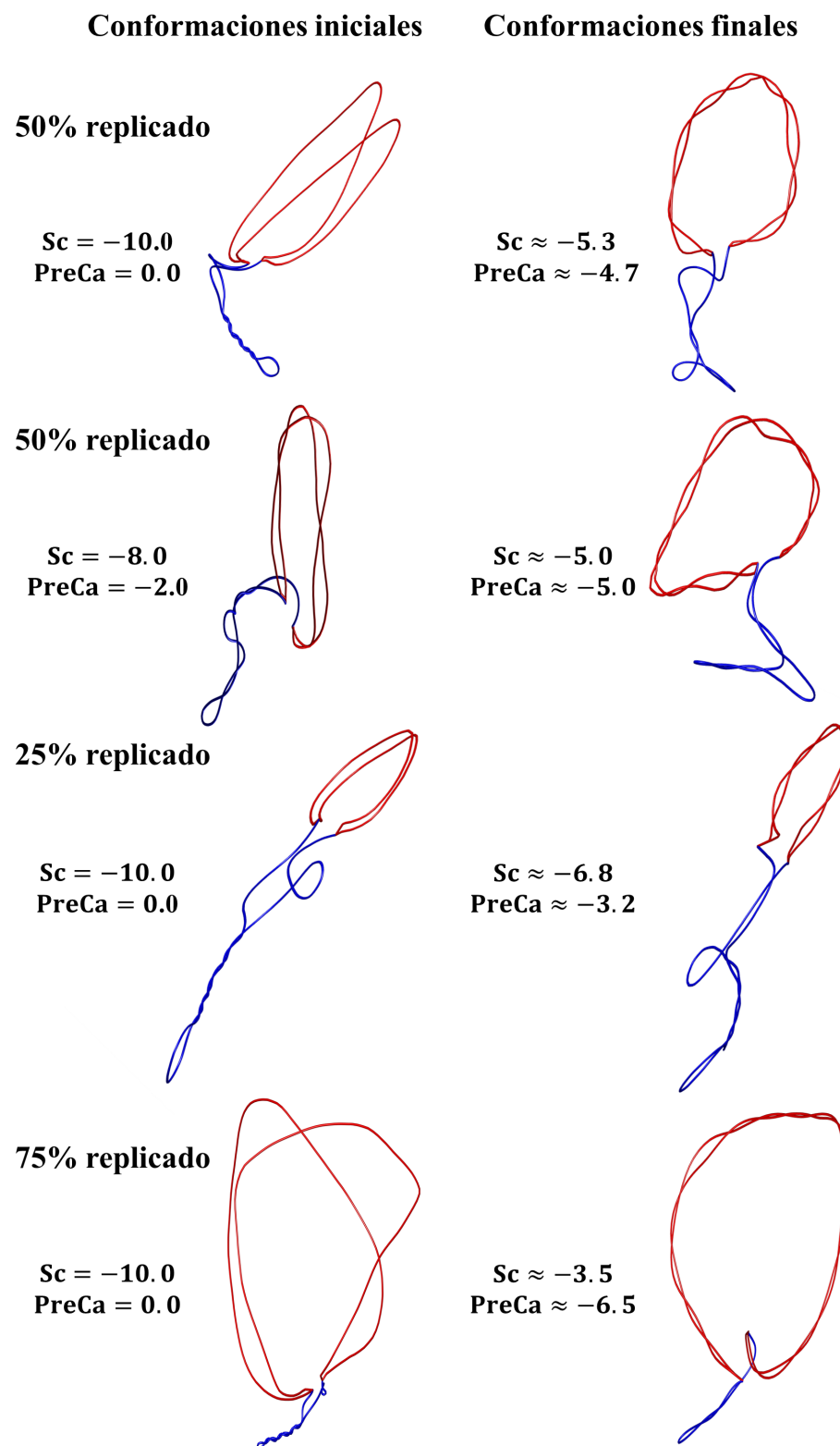


Figura 4.8: Conformaciones iniciales y finales de algunas simulaciones realizadas. Complejidad topológica de RIs con un $\Delta Lk = -10$ que han replicado el 25, 50 ó 75% de su DNA, antes y después de alcanzar el equilibrio termodinámico.

Los valores de energía inicial y final, así como el número de cruces de superenrollamiento y pre-encadenamiento iniciales y finales, se presentan en la Tabla 4.1. La comparación de las conformaciones de equilibrio y los valores de energía en condiciones de equilibrio termodinámico demostraron que la distribución del estrés torsional depende de la distancia entre las horquillas y el origen de replicación, es decir, cambia con el avance de las horquillas de replicación. En todos los casos en los que se alcanzó el equilibrio termodinámico, se observó que la suma de energía inicial de ambas regiones es mayor que la energía final del RI.

Tabla 4.1: Conformaciones y energías iniciales y finales de las simulaciones realizadas

Porcentaje de replicación	Conformación inicial		Energía inicial (Kcal/mol)			Energía final (Kcal/mol)			Conformación final	
	Sc	PreCa	Región no replicada	Región ya replicada	Total	Región no replicada	Región ya replicada	Total	Sc	PreCa
50%	10(-)	0	29,6	4,7	34,2	4,2	46,6	50,8	0,0	10,0(-)
	10(-)	0	27,2	4,7	31,9	13,7	13,7	27,3	5,3(-)	4,7(-)
	20(-)	0	135,2	8,0	143,2	40,7	40,7	81,4	11,2(-)	8,8(-)
	10(+)	0	25,6	5,7	31,3	11,4	11,4	22,8	5,4(+)	4,6(+)
	20(+)	0	120,5	6,2	126,7	35,8	35,8	71,6	10,8(+)	9,2(+)
	8(-)	2(-)	21,9	7,7	29,6	12,6	12,6	25,1	5,1(-)	4,9(-)
50% al revés	0	10(-)	2,9	50,2	53,1	35,6	12,6	48,2	10(-)	0,0
	0	10(-)	3,6	33,5	37,1	8,8	8,8	17,6	3,8(-)	6,2(-)
	2(-)	8(-)	6,7	24,1	30,8	13,6	13,6	27,2	5,3(-)	4,7(-)
25%	10(-)	0	21,1	6,1	27,2	12,5	12,5	25,0	6,8(-)	3,2(-)
75%	10(-)	0	157,4	5,3	162,7	15,6	15,6	31,3	3,5(-)	6,5(-)

Capítulo 5

Discusión

La importancia del estudio de las moléculas de DNA circular y los efectos que tiene la replicación en su topología radica en el hecho de que estas forman parte del genoma de varios organismos. En el caso de las bacterias, su DNA cromosómico y plásmídico se organiza en forma de moléculas circulares, y sus topoisomerasas son el blanco de algunos antibióticos [84]. El genoma de algunos virus como los causantes de la poliomielitis y el papiloma humano también consiste en DNA circular [4]. En organismos protistas, como los causantes del mal de chagas y la leishmaniasis, existen distintos tipos de DNA circular [8]. Si bien la gran mayoría del DNA de los eucariotas está contenido en largas cadenas lineales en los cromosomas, existen algunas excepciones como el DNA mitocondrial, los plásmidos en levaduras, los cloroplastos vegetales y los círculos extracromosómicos de DNA (eccDNAs). Los eccDNAs contribuyen a la variación genética en las células somáticas en organismos multicelulares y la evolución de los eucariotas unicelulares [85]. Recientemente, los eccDNAs se han relacionado con cáncer y resistencia a medicamentos [86], planteando nuevas interrogantes sobre su biogénesis, función y utilidad clínica.

Estudios anteriores demostraron que una vez finalizada la replicación del DNA circular, las cadenas nacientes recién sintetizadas están muy encadenadas entre sí [49][87][88]. También se demostró que los encadenados *in vivo* provienen de pre-encadenados formados durante la replicación [49][64][70][89]. A su vez, la formación de estos pre-encadenados es el resultado de la migración del estrés torsional desde la región no replicada a la región ya replicada provocada por la rotación de las horquillas de replicación mientras el RI intenta alcanzar un torque nulo [1][67][70].

En todos los casos estudiados hasta ahora *in vivo*, los signos de los cruces de encadenamiento entre cromátidas hermanas totalmente replicadas son positivos [88]. Como estos encadenados provienen de los pre-encadenados, los cruces de estos últimos también deberían ser positivos. Anteriormente se mencionó que el pre-encadenamiento

in vivo se forma como consecuencia de la migración del estrés torsional desde la región no replicada hacia la ya replicada, mediante la rotación de las horquillas de replicación. Por lo tanto, debemos concluir que los cruces de pre-encadenamiento (+) formados *in vivo* derivan del superenrollamiento (+) que se acumula transitoriamente inmediatamente por delante de las horquillas.

Un estudio realizado mediante mediciones directas de torque de Le et al. demostró que en el DNA desnudo (sin nucleosomas) el módulo de torsión de dos cadenas trenzadas es más bajo que el de una sola cadena, lo cual sugiere que el estrés torsional generado por la replicación se almacena principalmente en la región ya replicada. Por el contrario, se demostró que una sola fibra de cromatina (DNA unido a nucleosomas), como el DNA de la región no replicada, tiene un módulo de torsión más bajo que una fibra trenzada, como las cromátidas hermanas de la región ya replicada [90]. Estos resultados sugieren que en el caso de los organismos eucariotas el estrés torsional se almacena preferentemente en la región no replicada, mientras que en los procariotas sería más fácil introducir estrés torsional en la región ya replicada. Lo observado por Le et al. coincide con nuestros resultados experimentales que demuestran que en procariotas el estrés torsional se difunde continuamente de la región no replicada a la ya replicada durante la replicación del DNA [1]. Por otro lado, Schalbetter et al., demostraron que en eucariotas la difusión del estrés torsional tiene lugar hacia el final de proceso [91].

Tras el aislamiento del DNA en condiciones *in vitro*, una vez que se eliminan todas las proteínas, las horquillas de replicación de los RIs pueden girar libremente y el torque que genera este giro proviene del gradiente de la energía acumulada por el estrés torsional entre las regiones no replicada y ya replicada. Teniendo esto en cuenta, en la presente obra desarrollamos un modelo de Intermediario de Replicación basado en el modelo WLC para estudiar el intercambio entre superenrollamiento y pre-encadenamiento en moléculas parcialmente replicadas en condiciones *in vitro* e *in vivo*.

Este es el primer estudio del intercambio de estrés torsional entre las regiones no replicada y ya replicada en un Intermediario de Replicación de DNA utilizando un enfoque numérico. La implementación del modelo WLC utilizada en este estudio supone que, para estudiar el intercambio entre el superenrollamiento y pre-encadenamiento en un RI, debemos considerar los componentes de doblado y torsión de la energía elástica. La energía de doblado usualmente utilizada en el modelo WLC es normalmente representada como muestra la ecuación 2.23. Por otro lado, el WLC simple no incluye explícitamente un grado de libertad asociado a torsión. Existen diferentes enfoques para calcular el componente torsional de la energía, incluido el método *twistable worm – like chain* (TWLC)[92] o el modelo *helical worm – like chain*[93]. Dado que en esta obra no estamos interesados en efectos

locales de la dinámica de torsión, o en la falta de homogeneidad de la distribución de la torsión a lo largo de la cadena, la integral de todos los ángulos de torsión a lo largo de la cadena se puede utilizar para definir el componente torsional de la energía elástica [94]. En la presente tesis se utilizó la expresión de la energía elástica de torsión propuesta por Vologodskii et al. que, a pesar de su simplicidad, describe muchas propiedades termodinámicas del DNA [81]. Esta implementación del modelo WLC tiene varias suposiciones y aproximaciones, las cuales son: la aproximación del DNA como una cadena elástica isotrópica, por lo que se ignoran las variaciones locales en la estructura que dependen de la secuencia; no se consideran transiciones estructurales impulsadas por el superenrollamiento, como la introducción de cruciformes y formas Z-DNA; la suposición de que la rigidez torsional y de doblado son independientes; no se consideran los efectos del estrés de tensión por lo que esta implementación del modelo no cuenta con energía asociada al estiramiento; la aproximación de interacciones electrostáticas entre segmentos de DNA por cilindros cuyo diámetro efectivo explica los efectos de volumen excluidos [81]. A pesar de estas limitaciones, el procedimiento de simulación utilizado aquí reproduce las propiedades físicas del DNA superenrollado y encadenado [49][80][81], las cuales se corresponden respectivamente a las propiedades conformacionales de las regiones no replicada y ya replicada de un RI.

Para moléculas con las horquillas de replicación estancadas después de que hayan replicado la mitad del DNA, donde ambas regiones del RI simulado tienen la misma longitud, se encontró bastante similitud entre valores de energía de las regiones no replicada y ya replicada en el estado de equilibrio termodinámico. De igual manera, también se observó que en la conformación de equilibrio, la distribución del estrés torsional muestra una fuerte tendencia a ser del 50% del estrés torsional acumulado en la región no replicada en forma de cruces de superenrollamiento, mientras que la otra mitad termina albergada en la región ya replicada en forma de cruces de pre-encadenamiento.

Estos resultados mostraron mucha similitud entre sí tanto para moléculas con estrés torsional negativo o positivo, indicando que las propiedades termodinámicas y la forma en que se distribuye la deformación torsional son independientes del signo topológico para esta implementación del modelo WLC. Por otro lado, el estudio de la distribución de energía en función de la complejidad en RIs que alcanzaron el 25, 50 y 75% de la replicación del DNA demostró que, en el estado de equilibrio termodinámico, la distribución del estrés torsional depende de la progresión de las horquillas de replicación. Con estos resultados se puede concluir que el estrés torsional acumulado en las regiones no replicada y ya replicada depende de sus tamaños relativos, es decir, de cuanto ha progresado la replicación del DNA. En este sentido, nuestros resultados confirman lo propuesto por Ullsperger et al., que la distribución de equilibrio del ΔLk

entre regiones no replicada y ya replicada debería aproximarse al cociente entre sus longitudes [70].

En nuestras simulaciones numéricas, el gradiente de energía entre ambas regiones fue significativamente alto al principio, cuando todo el esfuerzo de torsión se acumuló en una región (no replicada o ya replicada). Una vez que se permitió la rotación de las horquillas de replicación para lograr el equilibrio del torque, la diferencia de energía elástica entre ambas regiones se redujo a medida que el esfuerzo de torsión migró de la región con mayor complejidad hacia la región más relajada. La tasa de cambio del gradiente de energía entre ambas regiones del RI fue más rápida al principio y disminuyó gradualmente a medida que el RI se acercaba al equilibrio termodinámico. El torque neto, consecuencia del estrés torsional local impuesto, provocó la rotación de las horquillas hasta alcanzar la conformación de equilibrio.

Una ventaja importante del enfoque de Metrópolis MC para estudiar las propiedades topológicas del DNA es que permite estudiar propiedades termodinámicas y conformacionales. En particular, en lo referente a la variación de entalpía, Vologodskii et al. la definieron en función de la energía elástica, cuyas componentes para nuestro modelo fueron calculadas con las ecuaciones 2.23 y 2.25. La variación de entalpía definida por Vologodskii et al. está dada como la diferencia entre la energía elástica de una conformación en estado de equilibrio termodinámico y la energía que tendría la molécula relajada (ΔLk nulo) [81]. Para el caso de simulación de RIs que alcanzaron el 50% de replicación la energía final en el estado de equilibrio termodinámico tiende a tener valores similares sin importar si las simulaciones partieron de diferentes condiciones iniciales respecto a la distribución del ΔLk . Como consecuencia directa de esto podemos afirmar que la diferencia de energía entre un RI relajado y uno con estrés torsional almacenado, o la variación de entalpía del RI, tiende a tener valores similares bajo la condición de que los RIs tengan la igual masa, misma proporción de tamaño entre ambas regiones y mismo ΔLk .

En el caso de RIs que alcanzaron el 50% de la replicación el contraste entre las energías de equilibrio entre los casos con $\Delta Lk = -10$ y $\Delta Lk = -20$, confirma que a mayor complejidad se tiene mayor energía elástica almacenada en el RI. Nuestros resultados sugieren que para pasar a un estado de menor complejidad las topoisomerasas liberan el exceso de energía almacenado en la molécula. Estudios experimentales llevados a cabo por nuestro grupo demostraron que los RIs digeridos *in vitro* con Topo IV se relajan completamente, dejando a los RIs relajados [1]. Por otro lado, bajo la acción *in vitro* de la DNA girasa, los RIs con ΔLk (-) deberían aumentar su complejidad hasta saturarse de cruces (-), mientras que los RIs con ΔLk (+) deberían ir perdiendo cruces (+) para luego convertirlos en cruces (-). Sin embargo, existen evidencias experimentales que sugieren

que la DNA girasa podría desencadenar cromátidas hermanas [72]. Por otro lado se ha demostrado que la DNA girasa es capaz de sustituir parcialmente a la Topo IV en las estirpes termosensibles parC o parE [95][96]. Estudios previos mostraron que pequeños cambios en la fuerza y el torque al que es sometido el DNA afectan la actividad de la DNA girasa mejorando su capacidad de introducir superenrollamiento (-), desencadenando cromátidas hermanas, o eliminando cruces LH (+) de superenrollamiento [97].

El método Metrópolis MC implementado en este estudio calcula la energía potencial elástica para cada región del RI, la cual depende de la distribución de la deformación elástica entre ambas regiones. En todos los casos en que el procedimiento de simulación se detuvo cuando ambas regiones alcanzaron el equilibrio termodinámico, se observa una diferencia en la energía elástica total entre la conformación inicial y final porque los perfiles de energía presentados no incluyen otros componentes de la energía total, como la energía cinética o la energía disipada en el medio. Tal diferencia fue más evidente en las moléculas que acumulan más estrés torsional, ($\Delta Lk = \pm 20$). En el caso de las moléculas replicadas al 75%, el tamaño de la región no replicada es significativamente menor y el efecto que tiene la contribución de la energía de doblado se incrementa. Como consecuencia, la energía inicial fue mayor que la de las moléculas replicadas hasta el 50%, que a su vez fue mayor que la energía inicial de los RIs que alcanzaron sólo el 25% de replicación del DNA. Independientemente del tamaño relativo de ambas regiones, una vez que las conformaciones llegaron al estado de equilibrio termodinámico presentaban valores de energía elástica similares. En el caso en que la energía elástica total final fue mayor que la energía total inicial (Figura 4.3B) la diferencia observada se debe al procedimiento de simulación que resulta en una conformación final que no podría haberse alcanzado a través de un proceso espontáneo. No obstante, es de interés el estudio de las propiedades energéticas en este estado.

Para el cálculo de la energía elástica de torsión es necesario conocer el valor del índice de superenrollamiento y nuestra implementación del algoritmo para calcular la integral de Gauss para el Wr (ecuación 3.5) efectivamente corre en tiempo $O(N^2)$. El proceso de cálculo de la energía elástica de torsión es el que requiere mayor tiempo computacional debido al cálculo de la integral doble de Gauss y por esto es importante como trabajo futuro la optimización de este algoritmo para mejorar el tiempo computacional en el cálculo del índice de superenrollamiento. La verificación del estado topológico de conformaciones pre-encadenadas se puede realizar fácilmente mediante la ecuación 3.1. Sin embargo, ese algoritmo se ejecuta en tiempo computacional $O(N^2)$ y lleva más tiempo de ejecución que el que calcula el Wr , ya que se deben recorrer dos curvas distintas para resolver esta integral doble de Gauss. El cálculo del grado de pre-encadenamiento utilizando un algoritmo basado en la ecuación 3.1 fue un medio de validación del método utilizado en esta obra, el cuál consistió en proyectar las ambas

curvas de la región ya replicada sobre un plano para contar el número de cruces entre ellas con su respectivo signo topológico. Por este motivo, el tiempo de ejecución del algoritmo de cálculo del número de pre-encadenamiento fue más bajo que el que utiliza la integral de Gauss.

Los movimientos de prueba para el método Metrópolis MC implementados en este trabajo fueron solamente las rotaciones rígidas descritas en la subsección 3.3.1, y con este movimiento ha sido posible reproducir varios fenómenos topológicos para el modelo utilizado. Sin embargo, en estudios posteriores podría implementarse un procedimiento de simulación que agregue nuevos tipos de movimientos de prueba. Tales movimientos de prueba pueden ser el movimiento de reptancia [81], o movimientos de reflexión, inversión o intercambio [98].

Otra aproximación teórica para simulaciones numéricas de Intermedios de Replicación podría probarse en trabajos futuros con un modelo que incluye más grados de libertad que los expuestos en el modelo implementado aquí, es decir, modelos en los que se tienen en cuenta las variaciones locales de los ángulos de torsión y los efectos del esfuerzo de tensión. Por su parte, el estudio de los efectos de torsión aportaría información sobre las contribuciones locales del esfuerzo de torsión en cada parte de la molécula replicante al torque neto dentro de ella.

Alternativamente, como una tarea futura, los métodos computacionales como las simulaciones de dinámica molecular podrían usarse para predecir la dinámica de la horquilla de replicación y la evolución de las moléculas simuladas a través del espacio de fases.

A todos estos tipos de análisis podría agregarse un estudio de simulación numérica sobre el movimiento de un RI dotado de carga eléctrica a través de un campo eléctrico dentro de un medio resistivo, es decir la movilidad electroforética. También mediante simulación se puede realizar un estudio en el cambio conformacional y termodinámico de los Intermediarios de Replicación bajo la acción directa de topoisomerasas.

Capítulo 6

Conclusiones

En esta obra se utilizó un abordaje basado en simulaciones computacionales para entender mejor la topología del DNA durante la replicación de moléculas de DNA circular parcialmente replicadas cuyas horquillas de replicación están detenidas tras haber logrado replicar el 25, 50 y 75% del DNA, siendo este el primer análisis de este tipo de moléculas realizado mediante simulación computacional. En las simulaciones numéricas se buscó caracterizar la relación entre las propiedades conformacionales, ligadas al intercambio entre superenrollamiento y pre-encadenamiento, y las propiedades termodinámicas de las moléculas de DNA parcialmente replicadas. Los resultados obtenidos demostraron que si bien la conformación de equilibrio de un Intermediario de Replicación de DNA circular es dinámica, ya que cambia a medida que avanzan las horquillas de replicación, el valor de su energía elástica total en el estado de equilibrio termodinámico es independiente del tamaño relativo de las regiones no replicada y ya replicada.

De acuerdo con los resultados obtenidos nuestras conclusiones son las siguientes.

- 1- Se desarrolló un modelo de moléculas de DNA circular parcialmente replicadas que permite estudiar el intercambio de estrés torsional entre las regiones no replicada y ya replicada mediante la rotación de ambas horquillas de replicación.
- 2- Se demostró que el gradiente de energía potencial elástica entre las regiones no replicada y ya replicada de un Intermediario de Replicación de DNA circular depende de la distribución del estrés torsional y el tamaño relativo de ambas regiones.
- 3- Se demostró que la distribución de equilibrio del estrés torsional entre las regiones no replicada y ya replicada de un Intermediario de Replicación de DNA circular es independiente del signo topológico.

4- Se demostró que la distribución de equilibrio del estrés torsional entre las regiones no replicada y ya replicada es independiente de la complejidad topológica del Intermediario de Replicación de DNA circular.

5- Se confirmó que la distribución de equilibrio del estrés torsional entre las regiones no replicada y ya replicada de un Intermediario de Replicación de DNA circular depende de la distancia entre las horquillas y el origen de replicación, es decir, está relacionada con el cociente entre el tamaño de ambas regiones.

REFERENCIAS BIBLIOGRÁFICAS

- [1] *Direct evidence for the formation of precatenanes during DNA replication.* Cebrian J., Castan A., Martinez V., Kadomatsu-Hermosa M.J., Parra C., Fernandez-Nestosa M.J., Schaerer C., Hernandez P., Krimer D.B. , Schwartzman J.B. 2015, J. Biol. Chem. 290, págs. 13725–13735.
- [2] *Molecular structure of nucleic acids.* Crick., J. D. Watson and F. H. C. 1953, Nature, 161, págs. 737–738.
- [3] *A topological view of replicon.* J. B. Schwartzman, Andrzej Strasiak. 2004, EMBO Reports, págs. 256-261.
- [4] *Small DNA circles as probes of DNA topology.* Bates A., Noy A., Piperakis M., Harris S., Maxwell A. 2013, Biochem. Soc. Trans. 41., págs. 565–570.
- [5] *Extrachromosomal Circular DNA of Tandemly Repeated Genomic Sequences in Drosophila.* Cohen, S., Yacobi, K., Segal, D. 2003, Genome Research 13 (6A), págs. 1133 - 1145.
- [6] *Identification of Autonomously Replicating Circular Subtelomeric Y' Elements in Saccharomyces cerevisiae.* Horowitz, H., Haber, J. E. 1985, Molecular and cellular biology 5 (9), págs. 56 - 64.
- [7] *Mapping and sequencing of structural variation from eight human genomes.* Kidd, J. M., Cooper, G. M.,. 2008, Nature. 453 (7191), págs. 56-64.
- [8] *Kinetoplast DNA Network: Evolution of an Improbable Structure.* Lukes J., Guilbride D., Voty J., Zíková A., Benne R, Englund P. 2002, EUKARYOTIC CELL, , págs. 495–502.
- [9] *Extrachromosomal circular DNA drives oncogenic genome remodeling in neuroblastoma.* Koche, R.P., Rodriguez-Fos, E., Helmsauer, K. 2020, Nat Genet 52 <https://doi.org/10.1038/s41588-019-0547-z>, págs. 29–34.

- [10] *Small Circular DNAs in Human Pathology*. Barreto S. C., Uppalapati M., Ray A. 2014, Malays J Med Sci. 21(3): , págs. 4-18.
- [11] Bates A., Maxwell A., *DNA Topology*. Great Clarendon Street, Oxford OX2 6DP Oxford Biocience, 2005.
- [12] *Closing the DNA replication cycle, from simple circular molecules to supercoiled and knotted DNA catenanes*. J. B. Schvartzman, Pablo Hernández, Dora B. Krimer, Julien Dorier, Andrzej Stasiak. 2019, Nucleic Acids Research, Vol. 47, No. 14, págs. 7182–7198.
- [13] *Phylogenomics of DNA topoisomerases: their origin and putative roles in the emergence of modern organisms*. Forterre P., Gadelles D. 2009, Nucleic Acids Res 37, págs. 679-692.
- [14] *Mapping the phase diagram of the writhe of DNA nanocircles using atomistic molecular dynamics simulations*. Harris S., Laughton C., Liverpool T. 2008, Nucleic Acids Research, Vol. 36, No. 1 , págs. 21–29.
- [15] Solomon, Berg, Martin, Villee. *Biología*. 3ra ed. México , Editorial Interamericana, Mc Graw – Hill, 1996.
- [16] *Sequence-dependent helical periodicity of DNA*. Rhodes D., Klug A. 1981, Nature, 292, págs. 378–380.
- [17] Bray, Bruce Alberts and Dennis. *Biología molecular de la célula*. s.l. , Ediciones Omega, 2010.
- [18] *Small DNA circles as probes of DNA topology*. Bates Andrew., Noy A., Piperakis M., Harris S., Maxwell A. 2013, Biochem. Soc. Trans. 41, págs. 565–570.
- [19] *DNA looping: the consequences and its control*. Saiz L., Vilar J. 2006, Current Opinion in Structural Biology (16), págs. 344–350.
- [20] Seifert, H. Threlfall, W. *A Textbook of Topology*. New York : Academic Press, 1980.
- [21] *The bacterial chromosome and its manner of replication as seen by autoradiography*. Cairns, J. 1963, J Mol Biol, 6, págs. 208–213.
- [22] A. Worcel, E. Burgi. 1972, J Mol Biol, 71, págs. 127–147.
- [23] *Interaction between DNA and an Escherichia coli protein*. Wang J. C. 1971, J Mol Biol, 55, págs. 523–533.
- [24] *Cellular roles of DNA topoisomerases, a molecular perspective*. Wang, J. C. 2002, Nat Rev Mol Cell Biol, 3, págs. 430–440.

- [25] *Combining high-resolution AFM with MD simulations shows that DNA supercoiling induces kinks and defects that enhance flexibility and recognition.* Pyne A., Agnes Noy A., Kavita Main K., Velasco-Berrelleza V., Piperakis M., Mitchenall L., Cugliandolo F., Beton J., Stevenson C., Hoogenboom B. W., Bates A. D., Maxwell A., Harris S. 2020, bioRxiv preprint doi: <https://doi.org/10.1101/863423>.
- [26] *Twisted circular form of polyoma viral DNA.* Vinograd, J., Lebowitz J., Radloff R., Watson R., P. Laipis. 1965, Proc. Natl. Acad. Sci., 53, págs. 1104–1111.
- [27] *DNA topology: fundamentals.* Mirkin, S. M. 2001, eLS, Wiley Online Library.
- [28] Sadiku M. *Elementos de electromagnetismo.* Mexico . Oxford University Press Mexico, 2003.
- [29] Gauss, Carl Friedrich. *Integral formula for linking number, in Zur mathematischen theorie der electrodynamische wirkungen.* Collected Works, Vol 5, Koniglichen Gesellschaft des Wissenschaften. Gottingen . s.n., 1833, pág. 605.
- [30] *Computation of writhe in modeling of supercoiled DNA.* K. Klenin, Jorg Langowski. 2000, Division Biophysics of Macromolecules, German Cancer Research Center, 3, págs. 430–440.
- [31] Bruce, H. Edwards Roland E. Larson, Robert Paul Hostetler. *Cálculo II.* s.l. : Mc Graw Hill, 2006.
- [32] *On the writhing number of a non-closed curve.* Starostin., E. L. 2005, Physical and Numerical Models in Knot Theory, págs. 525–545.
- [33] *L'integral de Gauss et l'analyse des noeuds tridimensionnels.* Calugareanu, G. 1959, Czechoslovak Math 11, págs. 588–625.
- [34] *Geometry of calugareanu's theorem.* M. R. Dennis, J. H. Hannay. 2005, URL <http://arxiv.org/abs/math-ph/0503012v2>.
- [35] *The writhe of open and closed curves.* M. A. Berger, C. Prior. 2006, UCL Mathematics Dept. J. Phys. A, 39, págs. 8321–8348.
- [36] *Supercoiled DNA.* Crick F., H. C. Bauer, W. R., J. H. Wihte. 1980, Sci Am, 243, págs. 100–118.
- [37] *Analysis of DNA topology of EBV minichromosomes in HEK 293 cells.* A. Castán, V. Fernández-Calleja, P. Hernández, Dora B. Krimer, J. B. Schwartzman, M. J. Fernández-Nestosa. 12, s.l., Plos One, 2017.
- [38] *Structure of plectonemically supercoiled DNA.* J.H. White, T.C. Boles, N.R. Cozzarelli. 1990, J Mol Biol, 213, págs. 931–951.

- [39] *DNA topoisomerases*. Gellert, M. 1981, Annu Rev Biochem, 50, págs. 879–910.
- [40] *DNA topoisomerases, enzymes that control dna conformation*. Vosberg, H. P. 1985, Curr Top Microbiol Immunol, 114, págs. 19–102.
- [41] *DNA topoisomerases*. J. C. Wang. 1985, Annu Rev Biochem, 54, págs. 665–697.
- [42] *Recent studies of DNA topoisomerases*. Wang, J. C. 1987, Biochim Biophys Acta, 909, págs. 1–9.
- [43] *DNA supercoiling in vivo*. L. Snyder, G. Giaever, J. C. Wang. 1988, Biophys Chem, 29, págs. 7–15.
- [44] *Supercoiling of intracellular DNA can occur in eukaryotic cells*. G. N. Giaever, J. C. Wang. 1988, Cell, 55, págs. 849–856.
- [45] *ATP-dependent DNA topoisomerase from the archaeobacterium sulfolobus acidocaldarius. relaxation of supercoiled DNA at high temperature*. D. M. Mirambeau, P. Forterre. 1984, J Mol Biol, 179, págs. 559–563.
- [46] *DNA topoisomerases*. Wang J. C. 1996, Annu Rev Biochem, 65, págs. 635–692.
- [47] *DNA gyrase action involves the introduction of transient double-strand breaks into DNA*. Mizuuchi, K., Fisher, L. M., O’Dea, M. H. & Gellert, M. 1980, Proc. Natl Acad. Sci. USA 77, págs. 1847–1851.
- [48] *DNA gyrase and the supercoiling of DNA*. Cozzarelli, N. R. 1980, Cience, 207, págs. 1104–1111.
- [49] *Interplay of DNA supercoiling and catenation during the segregation of sister duplexes*. María Luisa Martínez-Robles, Pablo Hernández Jorge B. Schvartzman Andrzej Stasiak, Guillaume Witz and Dora B. Krimer. 2009, Nucleic Acids Research, 37 (15), págs. 5126–5137.
- [50] *Terminal stages of sv40 DNA replication proceed via multiply intertwined catenated dimers*. O. Sundin, A. Varshavsky. 1980, Cell, 21, págs. 103–114.
- [51] *Catenation and knotting of duplex DNA by type I topoisomerases, a mechanistic parallel with type II topoisomerases*. P.O. Brown, N.R. Cozzarelli. 1981, Proc. Natl. Acad. Sci. USA 78, págs. 843–847.
- [52] *DNA transport by a type II topoisomerase, direct evidence for a two-gate mechanism*. J. M. Berger, S. C. Harrison, J. Roca, J. C. Wang. 1996, Proc Natl Acad Sci USA, 93, págs. 4057–4062.
- [53] *TypeII DNA topoisomerase genes*. Huang, W. M. 1994, Ad Pharmacol.

- [54] *New topoisomerase essential for chromosome segregation in E. Coli.* H. Suzuki, Nishimura Y. Imamura R. . Hiraga S. Kato. 1990, Cell, 63, págs. 393–404.
- [55] *Gene organization in the region containing a new gene involved in chromosome partition in E. Coli.* S. Hiraga, Nishimura Y. Yam M. H. Suzuki H. Kato. 1988, J Bacteriol, 170, págs. 3967–3977.
- [56] *Topoisomerase IV, not gyrase, decatenates products of site-specific recombination in Escherichia Coli.* Khodursky A. B., Zechiedrich, E. L. and N. R. Cozzarelli. 1997, Genes and Dev, 11, págs. 2580–2592.
- [57] *Topoisomerase IV alone unknots DNA in E. Coli.* Rahmati S. Deibler, R. W. and E. L. Zechiedrich. 2001, Genes and Dev, 15, págs. 748–761.
- [58] *Preferential relaxation of positively supercoiled DNA by E. coli topoisomerase IV in single-molecule and ensemble measurements.* Crisona N.J., Strick T.R., Bensimon D., Croquette V., Cozzarelli N.R.,. 2000, Genes Dev. 14, págs. 2881–2892.
- [59] *Chirality sensing by Escherichia coli topoisomerase IV and the mechanism of type II topoisomerases.* Stone M.D., Bryant Z., Crisona N.J., Smith S.B., Vologodskii A., Bustamante C., Cozzarelli, N.R.,. 2003, Proc. Natl. Acad. Sci. U.S.A. 100, págs. 8654–8659.
- [60] *The twist, writhe and overall shape of supercoiled DNA change during counterion induced transition from a loosely to a tightly interwound superhelix. Possible implications for DNA structure in vivo.* Bednar J., Furrer P, Stasiak A, Dubochet J, Egelman EH, Bates AD The twist,. 1994, J Mol Biol 235, págs. 825–847.
- [61] *Use of site-specific recombination as a probe of DNA structure and metabolism in vivo.* Bliska JB, Cozzarelli NR,. 1987, J Mol Biol 194, págs. 205–218.
- [62] *Replication of superhelical DNAs in vitro.* Mariani KJ, Minden JS, Parada C. 1986, Prog Nucleic Acids Res Mol Biol 33, págs. 111–140.
- [63] *Mechanisms of separation of the complementary strands of DNA during replication.* Alexandrov AI, Cozzarelli NR, Holmes VF, Khodursky AB, Peter BJ, Postow L, Rybenkov V, Vologodskii AV. 1999, Genetica 106, págs. 131–140.
- [64] *The structure of supercoiled intermediates in DNA replication.* Peter BJ, Ullsperger C, Hiasa H, Mariani KJ, Cozzarelli NR. 1988, Cell, 94, págs. 819–827.
- [65] *Transcription forms and remodels supercoiling domains unfolding large-scale chromatin structures.* Naughton, C., Avlonitis, N., Corless, S., Prendergast, J.G., Mati, I.K., Eijk, P.P., Cockroft, S.L., Bradley, M., Ylstra, B., and Gilbert, N. 2013, Nat. Struct. Mol. Biol. 20, págs. 387–395.

- [66] *Topological domains in mammalian genomes identified by analysis of chromatin interactions*. Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. 2012, *Nature* 485, págs. 376–380.
- [67] *Topoisomerases and the swivel problem*. In *Mechanistic Studies of DNA Replication and Genetic Recombination*. Champoux, J.J. and Been, M.D. 1980, Academic Press, New York, págs. 809–815.
- [68] *The Causes and Consequences of Topological Stress during DNA Replication*. Keszthelyi, A., Minchell, N.E., and Baxter, J. 2016, *Genes (Basel)* 7.
- [69] *DNA Catenation Reveals the Dynamics of DNA Topology During Replication*. A. Castán, P Hernández, DB Krimer, JB Schwartzman. 2018, *methods mol biol*, págs. 75–86.
- [70] *Unlinking of DNA by topoisomerases during DNA replication*. Ullsperger, C., Vologodskii, A.A., and Cozzarelli, N.R. 1995, In *Nucleic Acids and Molecular Biology*, D.M.J. Lilley, and F. Eckstein (Berlin, Springer-Verlag), págs. 115–142.
- [71] *Fork rotation and DNA precatenation are restricted during DNA replication to prevent chromosomal instability*. Schalbetter, S.A., Mansoubi, S., Chambers, A.L., Downs, J.A. and Baxter, J. 2015, *Proc. Natl. Acad. Sci. U.S.A.*, 112 *Proc. Natl. Acad. Sci. U.S.A.*, 112, págs. 4565–4570.
- [72] *Decatenating activity of Escherichia coli DNA gyrase and topoisomerases I and II during oriC and pBR322 DNA replication in vitro*. Hiasa H, Digate RJ, Mariani KJ. 1994, *J Biol Chem* 269, págs. 2093–2099.
- [73] *Knot what we thought before, the twisted story of replication*. Postow L, Peter BJ, Cozzarelli NB. 1999, *BioEssays* 21, págs. 805–808.
- [74] *Formation of knots in partially replicated DNA molecules*. Sogo JM, Stasiak A, Martínez-Robles ML, Krimer DB, Hernández P, Schwartzman JB. 1999, *Mol Biol* 286, págs. 637–643.
- [75] *Topological challenges to DNA replication, Conformations at the fork*. Peter BJ, Hardy CD, Cozzarelli NR, Postow L, Crisone NJ. 2001, *Proc Natl Acad Sci* 98 (3), págs. 8219–8226.
- [76] *Molecular dynamics simulations of DNA and its complexes*. Khalid S., Rodger P.M. 2004, *Progress in Reaction Kinetics and Mechanism*, 29 (3), págs. 167–186.
- [77] *Equation of state calculations by fast computing machines*. Rosenbluth M., Teller A., Metropolis N., Rosenbluth A. 1953, *The Journal of Chemical Physics* 21(6), págs. 1500–1505.

- [78] *Statistical-mechanics of supercoils and the torsional stiffness of the DNA double helix*. Anshelevich V.V., Lukashin A.V. Vologodskii, A.V. and M.D. Frank-Kamenetskii. 1979, Nature, 280, págs. 294–298.
- [79] *Monte carlo computation of the supercoiling energy, the sedimentation constant and the radius of gyration of unknotted and knotted circular DNA*. Bret, M. Le. 1980, Biopolymers, págs. 619–637.
- [80] *Torsional and bending rigidity of the double helix from data on small DNA rings*. Lukashin A. V., Anshelevich V. V. Vologodskii A. V. Frank-Kamenetskii, M. D. 1985, J. Biomol. Struct. Dynam, 2, págs. 1005–1012.
- [81] *Conformational and thermodynamic properties of supercoiled DNA*. Vologodskii A.V., Levene S.D., Klenin K.V. Frank-Kamenetskii M., and N.R. Cozzarelli. 1992, J. Mol. Biol, 227, págs. 1224–1243.
- [82] *Simulation of DNA Catenanes*. Vologodskii A., Rybenkov V. 2009, Phys Chem Chem Phys , págs. 10543–10552.
- [83] J. L. Movilla, Fernando Rajadell. *Termodinámica química*. Universidad Jaume. Av de Vicent Sos Baynat, s/n, 12071 Castelló de la Plana, Castellón, España , s.n., 2008.
- [84] *The antibiotic microcin B17 is a DNA gyrase poison: characterisation of the mode of inhibition*. Heddle, J.G., Blance, S.J., Zamble, D.B., Hollfelder, F., Miller, D.A., Wentzell, L.M., Walsh, C.T., and Maxwell, A. 2001, J. Mol. Biol. 307., págs. 1223–1234.
- [85] *Genome-wide Purification of Extrachromosomal Circular DNA from Eukaryotic Cells*. Møller H., Bojsen R., Tachibana C., Parsons L., Botstein D., Regenberg B. 2016, <https://www.jove.com/video/54239>.
- [86] *New discoveries of extrachromosomal circles of DNA in normal and tumor cells*. Paulsen T., Kumar P., Koseoglu M., Dutta A. 2018, Trends Genet. 34(4), págs. 270–278.
- [87] *Action at hooked or twisted-hooked DNA juxtapositions rationalize unlinking preference of type 2 topoisomerases*. Liu Z., Zechiedrich L., Chan H. S. 2010, Journal of molecular biology, 400, págs. 963–982.
- [88] *Arrest of segregation leads to accumulation of highly intertwined catenated dimers dissection of the final stages of SV40 DNA replication*. Sundin, A. Varshavsky,. 1981, Cell 25, págs. 659–669.

- [89] *The role of topoisomerase IV in partitioning bacterial replicons and the structure of catenated intermediates in DNA replication.* Adams D.E., Shekhtman E.M., Zechiedrich E.L., Schmid M.B., Cozarelli N.R., 1992, Cell, 71, págs. 277–288.
- [90] *Synergistic Coordination of Chromatin Torsional Mechanics and Topoisomerase Activity.* Tung T. Le, Xiang Gao, Seong Ha Park, Ryan P. Badman, James M. Berger, Michelle D. Wang. 2019, Cell 179, págs. 619–631.
- [91] *Fork rotation and DNA precatenation are restricted during DNA replication to prevent chromosomal instability.* S.A. Schalbetter, Mansoubi, S., Chambers, A.L., Downs, J.A., and Baxter, J. 2015, Proc Natl Acad Sci U S A 112, págs. 4565-4570.
- [92] *Twist-bend coupling and the statistical mechanics of the twistable worm-like chain model of DNA: perturbation theory and beyond.* Nomidis S K., Skoruppa E., Carlon E., Marko J. 2019, Phys Rev E. 99(3-1): 032414.
- [93] *Statistical mechanics of helical wormlike chain model.* Ya Liu., Toni Pérez, Wei Li, J. D. Gunton, and Amanda Green. 2011, J. Chem. Phys. 134, 065107.
- [94] *Evaluating changes of writhe in computer simulations of supercoiled DNA.* de Vries R, . 2005, The Journal of Chemical Physics 122, 064905.
- [95] *Purification and characterization of DNA topoisomerase IV in Escherichia coli.* Kato J., Suzuki H., Ikeda H. 1992, J Biol Chem. 267(36), págs. 25676-25684.
- [96] *Specificity and Redundancy in the Function of DNA Topoisomerases.* Cebrián J., Martínez V., Hernández P., Krimer D. B., Fernández-Nestosa M-J., Schwartzman J. B. 2020. Submitted article.
- [97] *Multiple modes of Escherichia coli DNA gyrase activity revealed by force and torque.* Nollmann, M., Stone, M.D., Bryant, Z., Gore, J., Crisona, N.J., Hong, S.C., Mitelheiser, S., Maxwell, A., Bustamante, C., and Cozzarelli, N.R. 2007, Nat Struct Mol Biol 14, , págs. 264-271.
- [98] *Efficient chain moves for Monte Carlo simulations of a wormlike DNA model: Excluded volume, supercoils, site juxtapositions, knots, and comparisons with random-flight and lattice models.* Zhirong Liua., Hue Sun Chanb.,. 2008, THE JOURNAL OF CHEMICAL PHYSICS 128.

Anexos

A.1 Algoritmos de base

A.1.1 Polígonos iniciales

Para representar las tres regiones que componen un Intermediario de Replicación se crean tres curvas, $r1$ para representar la región no replicada, $r2$ y $r3$ para la region ya replicada. Como parámetros iniciales se requiere la longitud de la molécula L expresada en pb , el número de segmentos N y el porcentaje de replicación p .

Algorithm 1 CreadorRI

Require: L longitud de la molécula en pb , N número de vertices o segmentos del polígono, p porcentaje de replicación, d separación inicial entre las horquillas de replicación

Ensure: Vectores de coordenadas $x1, y1, z1$ de cada vertice del polígono $r1$ regular de radio $R1$. Vectores $x2, y2, z2$ de cada vertice del polígono $r2$ regular de radio $R2$. Vectores $x3, y3, z3$ de cada vertice del polígono $r3$ regular de radio $R3$.

$Nrep = Lp$ Número de segmentos de la región ya replicada

$NNorep = L - Lrep$ Número de segmentos de la región no replicada

$Lrep = Lp$ Longitud de la region ya replicada

$LNorep = L - Lrep$ Longitud de la región no replicada

$R1 = LNorep/2Pi$ Radio de la curva $r1$

$R2 = Lrep/2Pi$ $R3 = R2$ Radio de las curvas $r2$ y $r3$

$D = \sqrt{R1^2 - d^2/4} + \sqrt{R2^2 - d^2/4}$;

for $i = 1$ hasta $NNorep$ **do**

$x1(i) = R1\cos[2 * Pi * i/NNorep] - D$

$y1(i) = R1\sen[2 * Pi * i/NNorep]$

$z1(i) = 0$

end for

for $i = 1$ hasta $Nrep$ **do**

$x2(i) = R2\cos[2 * Pi * i/Nrep]$

$y2(i) = R2\sen[2 * Pi * i/Nrep]$

$z2(i) = l$

$x3(i) = R3\cos[2 * Pi * i/Nrep]$

$y3(i) = R3\sen[2 * Pi * i/Nrep]$

$z3(i) = -l$

end for

return $r1, r2, r3$

A.1.2 Cálculo del *Writhe*

Este algoritmo calcula el índice de superenrollamiento de una curva.

Algorithm 2 Writhe

Require: N número de segmentos de la curva
 Vector de coordenadas x, y, z de la curva
Ensure: Writhe Wr de la curva
 $Wr = 0$ Inicializa a cero el valor de Wr
for $i = 1$ hasta $N - 1$ **do**
 $b_i = [x(i+1) - x(i), y(i+1) - y(i), z(i+1) - z(i)]$
 for $j = i + 1$ hasta N **do**
 $b_j = [x(j+1) - x(j), y(j+1) - y(j), z(j+1) - z(j)]$
 $R_{ji} = [x(j) - x(i), y(j) - y(i), z(j) - z(i)]$
 $Mixto = \text{Productovectorial}_{Mixto}(b_i, b_j, R_{ji})$
 $Wr = Wr + \frac{Mixto}{\|R_{ji}\|^{\frac{3}{2}}}$
 end for
end for
 $Wr = -\frac{Wr}{2\pi}$
return Wr

A.1.3 Rotación rígida

Este algoritmo efectua la rotación de todos los puntos contenidos entre los puntos fijos A y B en torno a un eje que pasa por ellos.

Algorithm 3 RotacionRigida

Require: I, F índice de dos puntos de la curva
 Vector de coordenadas x, y, z de la curva
 Ángulo aleatorio φ
Ensure: Vector de coordenadas x, y, z entre los índices I y F de la curva
for $j = I + 1$ hasta $F - 1$ **do**
 $P = [x(j), y(j), z(j)];$
 $A = [x(I), y(I), z(I)]; pA = P - A;$
 $B = [x(F), y(F), z(F)];$
 $\mathbf{n} = (B - A) / \text{norm}(B - A);$
 $P_0 = [\mathbf{n} \cdot (P - A)]\mathbf{n} + \cos(\varphi)[(P - A) - [\mathbf{n} \cdot (P - A)]\mathbf{n}] \pm \sin(\varphi)[\mathbf{n} \times (P - A)]$
 $x(j) = P_0(1); y(j) = P_0(2); z(j) = P_0(3);$
end for
return x, y, z

A.1.4 Rotación de dos curvas

Este algoritmo efectua la rotación entre dos curvas para representar la rotación las entre cadenas de la región ya replicada.

Algorithm 4 RotacionHorquillas

Require: $r1$, $r2$ vectores de coordenadas de las dos curvas

Ángulo de rotación φ

Ensure: Vector de coordenadas $r1$, $r2$ de las dos curvas

$PM = (r1 + r2)/2$; punto medio entre las dos curvas

for $j = 1$ hasta n **do**

$Phi = (j - 1) * \varphi / (2 * n)$;

$I = j$;

$F = j + 1$;

 rotación de la curva $r1$

$A = PM_I$; $pA = r1_j - A$;

$B = PM_F$;

$\mathbf{n} = (B - A) / norm(B - A)$;

$r1_j = [\mathbf{n} \cdot (r1_j - A)]\mathbf{n} + cos(Phi)[(r1_j - A) - [\mathbf{n} \cdot (P - A)]\mathbf{n}] \pm sin(Phi)[\mathbf{n} \times (r1_j - A)]$

 rotación de la curva $r2$

$A = PM_I$; $pA = r2_j - A$;

$B = PM_F$;

$\mathbf{n} = (B - A) / norm(B - A)$;

$r2_j = [\mathbf{n} \cdot (r2_j - A)]\mathbf{n} + cos(Phi)[(r2_j - A) - [\mathbf{n} \cdot (P - A)]\mathbf{n}] \pm sin(Phi)[\mathbf{n} \times (r2_j - A)]$

end for

return $r1$, $r2$

A.1.5 Verificación del diámetro efectivo

Este algoritmo verifica si dos puntos cualquiera de la curva C están a una distancia d mayor que el diámetro efectivo $d_{efectivo}$.

Algorithm 5 DiametroEfectivo

Require: Vector de coordenadas x, y, z de la curva

valor del diámetro efectivo $d_{efectivo}$

Ensure: *FALSE* si no hay ninguna violación del $d_{efectivo}$, *TRUE* caso contrario.

$count = 0$ Inicializa a cero un contador

$N = sizeOf(x)$;

for $i = 1$ hasta $N - 1$ **do**

$u = (x(i + 1) - x(i), y(i + 1) - y(i), z(i + 1) - z(i))$ vector u

for $j = i + 1$ hasta N **do**

$v = (x(j + 1) - x(j), y(j + 1) - y(j), z(j + 1) - z(j))$ vector v

$w = (x(j) - x(i), y(j) - y(i), z(j) - z(i))$ vector w

$d = \frac{Producto_{vectorial_Mixto}(w, u, v)}{\|Producto_{vectorial}(u, v)\|}$ calcula la distancia entre dos rectas

$l = n \times \|(x(i) - x(j)), (y(i) - y(j)), (z(i) - z(j))\|$ calcula la distancia entre dos puntos, n depende de la cantidad de segmentos y la longitud de la molécula modelada

if $d \leq d_{efectivo}$ and $l < d_{efectivo}$ **then**

$count = count + 1$

end if

end for

end for

return $count$

A.2 Simulación del superenrollamiento

Algorithm 6 Superenrollamiento

Require: Número de segmentos N de la curva, su ΔLk , cantidad $Iter$ de movimientos de prueba.

vectores x , y y z que describen la curva modelada

T : Temperatura absoluta; k_b : Constante de Boltzmann; C : Constante de torsión; a :

Constante de doblado; L : Longitud de la cadena en pb

$Min = \infty$ Inicializa la energía a un valor muy grande Min

for $l = 1$ hasta $Iter$ **do**

if $l = 1$ **then**

$x = X_{aceptado}$, $y = Y_{aceptado}$, $z = Z_{aceptado}$

end if

I, F : índices aleatorios entre 1 y N para una rotación rígida;

φ : ángulo aleatorio.

RotacionRigida(x, y, z, I, F, φ) : Rotación de los puntos entre I y F

DiametroEfectivo($x, y, z, di - ef$)

Eb=EnergiaDoblado(x, y, z): El valor de la energía de doblado de la conformación

Wr = Writhe(x, y, z): Calcula el writhe de la curva

Et = $\frac{2\pi^2}{L}(\Delta Lk - Wr)^2$: La energía de torsión de la conformación

E=Eb+Et

$Prob = e^{-\frac{(E-Min)}{k_b T}}$

$\Delta = \text{PolinomioAlexander}(x, y, z)$: Verifica la topología de la conformación

if $\Delta \neq 1$ **then**

$E = \infty$

end if

if $E > Min$ **then**

if $Prob > ProbabilidadMinimaAceptada$ **then**

$Min = E$

end if

end if

if $E \leq Min$ **then**

$Min = E$

$X_{aceptado} = x, Y_{aceptado} = y, Z_{aceptado} = z,$

end if

end for

return $X_{aceptado}, Y_{aceptado}, Z_{aceptado}$

A.3 Simulación del pre-encadenamiento

Algorithm 7 Pre-encadenamiento: parte 1

Require: Número de segmentos N de cada curva, su $PreCa$, y la cantidad $Iter$ de movimientos de prueba. Vector de coordenadas x, y, z de una curva
 Vector de coordenadas x', y', z' de la otra curva
 T : Temperatura absoluta; k_b : Constante de Boltzman; C : Constante de rigidez de torsión; a : Constante de rigidez de doblado; L : Longitud de cada cadena en pb
 $Min = \infty$; $Min' = \infty$ Inicializa las energías a un valor muy grande Min
for $l = 1$ hasta $Iter$ **do**
 if $l = 1$ **then**
 $x = X_{aceptado}$, $y = Y_{aceptado}$, $z = Z_{aceptado}$; $x' = X'_{aceptado}$, $y' = Y'_{aceptado}$, $z' = Z'_{aceptado}$
 end if
 I, F, I', F' : índices aleatorios entre 1 y N para una rotación rígida
 φ, φ' : ángulo aleatorio de rotación para una rotación rígida
 RotacionRigida(x, y, z, I, F, φ) : Hace la rotación rígida entre los puntos de la curva comprendido entre I y F
 RotacionRigida($x', y', z', I', F', \varphi'$) : Hace la rotación rígida entre los puntos de la curva comprendido entre I' y F'
 DiametroEfectivo($x, y, z, di - ef$): Asigna un valor alto de energía si encuentra dos puntos con distancia menor al diámetro efectivo di-ef
 DiametroEfectivo($x', y', z', di - ef$)
 DiametroEfectivo2($x, y, z, x', y', z', di - ef$): Asigna un valor alto de energía si encuentra dos puntos entre ambas curvas con distancia menor al diámetro efectivo di-ef
 Eb=EnergiaDoblado(x, y, z): Devuelve el valor de la energía de doblado total de la conformación
 Eb'=EnergiaDoblado(x', y', z')
 E=Eb
 E'=Eb'
end for

Algorithm 8 Pre-encadenamiento: parte 2

```

for  $l = 1$  hasta  $Iteraciones$  do
   $Prob = e^{-\frac{(E - Min)}{k_b T}}$ 
   $\Delta = \text{PolinomioAlexander}(x, y, z)$  : Verifica el estado topológico de la
  conformación
  if  $\Delta \neq 1$  then
     $E = \infty$ 
  end if
   $Prob' = e^{-\frac{(E' - Min')}{k_b T}}$ 
   $\Delta' = \text{PolinomioAlexander}(x', y', z')$  : Verifica el estado topológico de la
  conformación
  if  $\Delta' \neq 1$  then
     $E' = \infty$ 
  end if
   $\text{VerificaPreEncadenamiento}(x, y, z, x', y', z', preCa)$  : Devuelve energías
  infinitas si las curvas tienen un pre-encadenamiento distinto de  $preCa$ , si no, no
  hace nada.
  if  $E > Min$  then
    if  $Prob > ProbabilidadMinimaAceptada$  then
       $Min = E$ 
    end if
  end if
  if  $E \leq Min$  then
     $Min = E$ 
     $X_{aceptado} = x, Y_{aceptado} = y, Z_{aceptado} = z,$ 
  end if
  if  $E' > Min'$  then
    if  $Prob' > ProbabilidadMinimaAceptada$  then
       $Min' = E'$ 
    end if
  end if
  if  $E' \leq Min'$  then
     $Min' = E'$ 
     $X'_{aceptado} = x', Y'_{aceptado} = y', Z'_{aceptado} = z',$ 
  end if
end for
return  $X_{aceptado}, Y_{aceptado}, Z_{aceptado}, X'_{aceptado}, Y'_{aceptado}, Z'_{aceptado}$ 

```

A.4 Cambio de conformación de un RI

Este algoritmo produce el cambio de conformación de un RI, redistribuyendo el ΔLk entre ambas regiones del RI. .

Algorithm 9 Cambio de conformación de un RI

Require: N número de segmentos del RI sin replicar

L longitud de la molécula

p porcentaje de replicación

T número total de vueltas que realizarán las horquillas

S número total de pasos

ΔLk índice de ligamiento de la molécula

Ensure: $r1, r2, r3, ERep, ENoRep$.

$(r1, r2, r3) = \text{CreadorRI}(N, L, p)$; se crea la conformación inicial

$\varphi = 2 * PI * T/S$ ángulo de rotación de las horquillas de replicación en cada paso

for $i = 1$ hasta S **do**

$(r1, ENoRep(i)) = \text{superenrollamiento}(r1, \Delta Lk - i \times T/S)$; aplica el método Metrópolis MC a la región no replicada. Cambia el grado de superenrollamiento en la región no replicada una magnitud T/S

$(r2, r3, ERep(i)) = \text{pre-encadenamiento}(r2, r3, preCa)$; aplica el método Metrópolis MC a la región ya replicada

$(r2, r3) = \text{RotacionHorquillas}(r2, r3, \varphi)$; gira las curvas $r2$ y $r3$ entre sí un ángulo φ . Cambia el grado de pre-encadenamiento en la región ya replicada una magnitud T/S

$preCa = \text{calculaCa}(r2, r3)$; calcula el nuevo valor del pre-encadenamiento entre $r2$ y $r3$

end for

return $r1, r2, r3, ERep, ENoRep$

A.5 Publicaciones originales derivadas de esta tesis doctoral

Distribution of Torsional Stress Between the Un-replicated and Replicated Regions in Partially Replicated Molecules. **Victor Martínez**, Christian Schaerer, Pablo Hernández, Dora B. Krimer, Jorge B. Schwartzman, María José Fernández-Nestosá. Journal of Biomolecular Structure & Dynamics. 2020 Apr 10; 1-12.

Topology of Partially Replicated DNA Molecules. **Victor Martínez**, Christian Schaerer, Pablo Hernández, Dora B. Krimer, Jorge B. Schwartzman, María José Fernández-Nestosá. Proceeding Series of the Brazilian Society of Computational and Applied Mathematics, v. 6, n. 2, 2018.



Distribution of torsional stress between the un-replicated and replicated regions in partially replicated molecules

Víctor Martínez, Christian Schaerer, Pablo Hernández, Dora B. Krimer, Jorge B. Schwartzman & María-José Fernández-Nestosa

To cite this article: Víctor Martínez, Christian Schaerer, Pablo Hernández, Dora B. Krimer, Jorge B. Schwartzman & María-José Fernández-Nestosa (2020): Distribution of torsional stress between the un-replicated and replicated regions in partially replicated molecules, Journal of Biomolecular Structure and Dynamics, DOI: [10.1080/07391102.2020.1751294](https://doi.org/10.1080/07391102.2020.1751294)

To link to this article: <https://doi.org/10.1080/07391102.2020.1751294>



View supplementary material [↗](#)



Accepted author version posted online: 02 Apr 2020.
Published online: 10 Apr 2020.



Submit your article to this journal [↗](#)



Article views: 11



View related articles [↗](#)



View Crossmark data [↗](#)



Distribution of torsional stress between the un-replicated and replicated regions in partially replicated molecules

Víctor Martínez^a, Christian Schaerer^a, Pablo Hernández^b, Dora B. Krimer^b, Jorge B. Schwartzman^b and María-José Fernández-Nestosa^a

^aPolytechnic School, National University of Asunción, San Lorenzo, Paraguay; ^bDepartment of Cellular and Molecular Biology, Centro de Investigaciones Biológicas (CSIC), Madrid, Spain

Communicated by Ramaswamy H. Sarma

ABSTRACT

DNA topology changes continuously as replication proceeds. Unwinding of the DNA duplex by helicases is favored by negative supercoiling but it causes the progressive accumulation of positive supercoiling ahead of the fork. This torsional stress must be removed for the fork to keep advancing. Elimination of this positive torsional stress may be accomplished by topoisomerases acting solely ahead of the fork or simultaneously in the un-replicated and replicated regions after diffusion of some positive torsional strain from the un-replicated to the replicated regions by swivelling of the replication forks. In any case, once replication is completed fully replicated molecules are known to be heavily catenated and this catenation derives from pre-catenanes formed during replication. Although there is still controversy as to whether fork swivelling redistributes this positive torsional stress continuously or only as termination approaches, the forces that cause fork rotation and the generation of pre-catenanes are still poorly characterized. Here we used a numerical simulation, based on the worm-like chain model and the Metropolis Monte Carlo method, to study the interchange of supercoiling and pre-catenation in a naked circular DNA molecule of 4,440 bp partially replicated *in vivo* and *in vitro*. We propose that a dynamic gradient of torsional stress between the un-replicated and replicated regions drives fork swivelling allowing the interchange of supercoiling and pre-catenation.

ARTICLE HISTORY

Received 14 January 2020
Accepted 18 March 2020

KEYWORDS

DNA topology; replication; supercoiling; pre-catenation
metropolis Monte Carlo

Introduction

DNA replication represents an important topological challenge. Due to the right-handed (RH) intertwined nature of the B-DNA double-helix, unwinding by DNA helicases causes a compensatory overwinding of the duplex ahead of the replication fork and this overwinding leads to positive (+) supercoiling (Kornberg & Baker, 1992). In bacterial cells negative (-) supercoiling facilitates the parental strands separation that is needed for the advance of DNA replication (Funnell et al., 1986; Marians et al., 1986; Martínez-Robles et al., 2009). The progressive accumulation of torsional stress in the form of (+) supercoiling immediately ahead of the fork would prevent its progression. During replication, elimination of this (+) supercoiling is accomplished by DNA topoisomerases acting solely ahead of the fork or simultaneously in the un-replicated and replicated regions after diffusion of the (+) supercoiling from the un-replicated to the replicated region by swivelling of the replication fork. In this way the un-replicated region remains always (-) supercoiled. This latter pathway, originally proposed by Champoux and Been (Champoux & Been, 1980), reduces the torsional stress immediately ahead of the replication fork at the expense of generating

(+) intertwined pre-catenanes in the replicated region (Champoux & Been, 1980; Ullsperger et al., 1995). It is important to notice that once replication is completed, fully replicated sister duplexes end-up (+) intertwined (Martínez-Robles et al., 2009; Sundin & Varshavsky, 1980, 1981).

Whether or not replication forks are able to rotate as they advance *in vivo* is still a matter of controversy (Cebrian et al., 2015; Mariezcurrena & Uhlmann, 2017; Schalbeter et al., 2015). In any case, after DNA isolation once all proteins are removed, *in vitro* the torsional stress of the un-replicated and replicated regions of a replication intermediate (RI) must equilibrate and this could only be accomplished by fork rotation (Schvartzman et al., 2019; Schvartzman & Stasiak, 2004). More specifically, in a free state, the system formed by both regions of the RI reaches thermodynamic equilibrium if the molecule is in mechanical equilibria, which in turn is achieved when the torque is balanced in the whole molecule. Winding or unwinding of the un-replicated portion of the plasmid applies torque to the replication forks, while braiding or unbraiding the replicated portion of DNA also applies torque to the forks but in the opposite direction, so that the net torque is cancelled in the equilibrium state.

CONTACT Jorge B. Schwartzman  schvartzman@cib.csic.es  Department of Cellular and Molecular Biology, Centro de Investigaciones Biológicas (CSIC), Ramiro de Maeztu 9, Madrid 28040, Spain; María-José Fernández-Nestosa  mjfernandez@pol.una.py  Polytechnic School, National University of Asunción, P.O. Box 2111, San Lorenzo, Paraguay.

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/07391102.2020.1751294>.

© 2020 Informa UK Limited, trading as Taylor & Francis Group

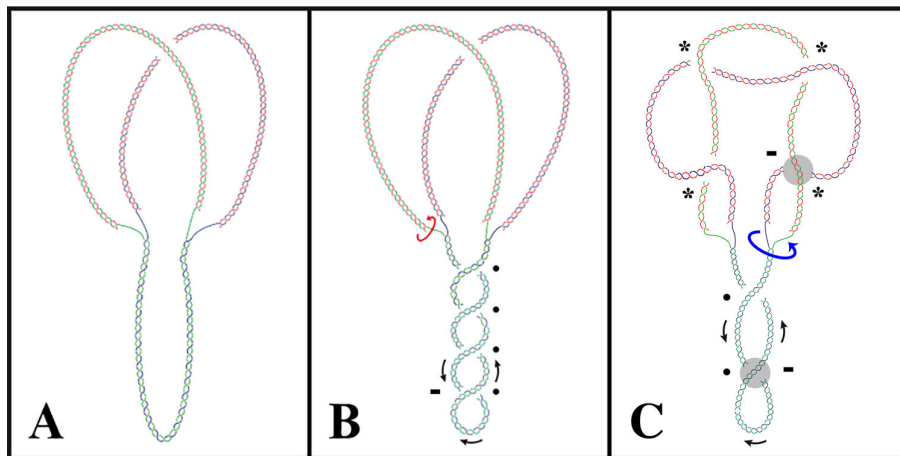


Figure 1. Cartoons illustrating the topology of partially replicated DNA molecules. (A): Fully relaxed replication intermediate. (B): Negatively supercoiled replication intermediate showing four (-) crossings in the un-replicated region (indicated by dots) and no pre-catenation. (C): Covalently closed replication intermediate showing two (-) crossings in the un-replicated region (indicated by dots) and four (-) pre-catenane crossings (indicated by asterisks) in the replicated one. Note that the migration of each (-) crossing from the un-replicated region to the replicated one leads to two (-) crossings. Parental DNA duplex is depicted in blue and green whereas the sister duplexes of partially replicated molecules are depicted in blue and red and green and red, respectively, indicating that they are the natural products of replication where the newly synthesized strands are depicted in red.

Another consequence of mechanical equilibrium is that the potential energy gradient between the two regions must be null. Torsional stress generated by replisome progression should partition ahead of and behind the fork to maintain a balance of torque (Le et al., 2019). An important aspect to be considered, given the particular conformation of RIs, is the topological sign in the distribution of torsional stress between the replicated and un-replicated regions. The molecule represented in Figure 1A corresponds to a half-replicated fully relaxed form. Figure 1B represents a half-replicated molecule with (-) supercoiling in the un-replicated region and no pre-catenation in the replicated one. This is how half-replicated molecules would look *in vivo* if fork rotation was prevented except at termination (Schalbetter et al., 2015; Sundin & Varshavsky, 1980, 1981). After DNA isolation, fork rotation would allow redistribution of the torsional strain between the un-replicated and replicated regions. Once equilibrium is achieved, the supercoils in the un-replicated region would remain (-) and the sister duplexes in the replicated region would wind around each other in a (-) manner, too (Figure 1C). Note that the diffusion of each (-) supercoil crossing from the unreplicated region to the replicated one generates two (-) pre-catenane crossings (Champoux & Been, 1980; Schwartzman & Stasiak, 2004; Ullsperger et al., 1995).

Most studies on DNA topology are performed on bacterial circular plasmids that are closed topological domains. During replication these small circular DNA molecules remain (-) supercoiled. Understanding the interplay between supercoiling and pre-catenation in partially replicated molecules is hindered, though, by the fact that DNA conformation is disturbed during sample preparation. On the other hand, several studies convincingly proved that computer simulation can give a very accurate description of the DNA properties that are difficult to address experimentally (Martinez-Robles et al., 2009; Rawdon et al., 2016; Vologodskii, 2011; Vologodskii & Rybenkov, 2009; Vologodskii & Cozzarelli, 1993;

Witz & Stasiak, 2010). For all these reasons, here we used the worm-like chain (WLC) model in combination with Metropolis Monte Carlo (MC) numerical simulation to study the dynamics of DNA topology in partially replicated plasmid DNA molecules, once these RIs are isolated *in vitro*. In the WLC model the DNA is represented as a chain of rigid cylinders of equal length, whose conformation can be described as a simple set of parameters: bending elasticity, torsional rigidity and effective double-helix diameter (Racko et al., 2017; Vologodskii, 1992; Vologodskii & Rybenkov, 2009; Vologodskii & Cozzarelli, 1993).

The topological and energetic characterization of circular DNA plasmids depends on the Linking number (Lk), which is an integer that measures the winding of the two strands of the double helix around each other and is unaffected by changes in the conformation of the DNA, as long as these do not involve the breakage of one or both strands. Lk is a topological invariant established by the Calugareanu-White-Fuller relation, $Lk = Wr + Tw$ (Fuller, 1971), where Tw (Twist) determines the degree of torsional deformation of the double helix and Wr (Writhe) is a number that describes the three-dimensional folding of the molecule (Klenin & Langowski, 2000). The torsional stress in a topologically closed DNA molecule can be expressed in terms of the linking number difference, $\Delta Lk = Lk - Lk_0$, where Lk_0 is the Lk for relaxed DNA. The topology of partially replicated plasmids also depends on the pre-catenation number that describes the interlinking between the two sister duplexes and equals half of the signed sum of intermolecular nodes (Sundin & Varshavsky, 1980, 1981). For replication intermediates isolated *in vitro* ΔLk should be distributed both ahead and behind the replication fork (Peter et al., 1998; Schwartzman et al., 2019; Schwartzman et al., 2013; Ullsperger et al., 1995). Multiple catenated DNA duplexes result after nicking and religation (Vologodskii & Cozzarelli, 1993; Wasserman et al., 1988). The ΔLk observed is a result of the

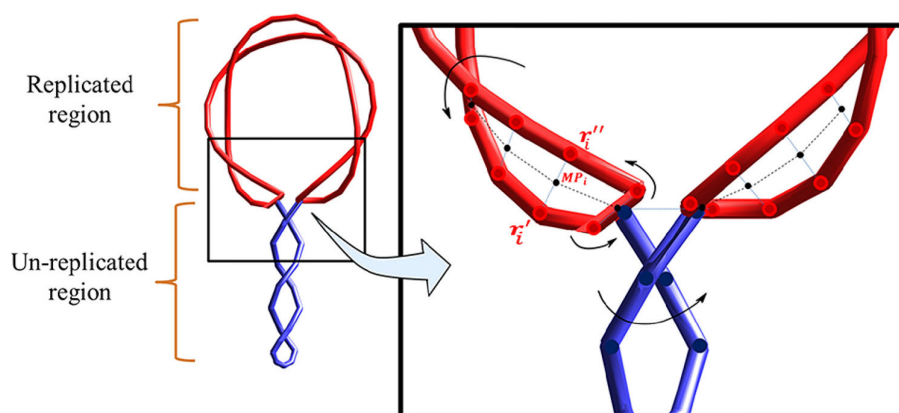


Figure 2. Simulation of fork rotation in partially replicated DNA molecules. Modeled replication intermediate (RI). Newly replicated sister DNA molecules multiply interlinked with (-) pre-catenane crossings (in red) and the un-replicated supercoiled region with (-) plectonemes (in blue). The mechanism to model fork rotation is shown in the inset figure. Opposite points of the replicated curves, r'_i and r''_i , located on each curve are rotated with respect to their midpoint MP_i over the same angle. Braiding or unbraiding of the replicated curves results from the contribution of rotations of all the pair of points of these curves. The topology of the un-replicated region is modified only by the effect of Metropolis MC method.

writhe induced by catenation. When catenated DNA molecules are supercoiled their extent of supercoiling is expressed by the effective ΔLk (ΔLk_e), which is defined as the difference between the actual Lk and the equilibrium linking number for the given catenated rings or knotted DNA molecule (Burnier et al., 2008; Schwartzman et al., 2019). In the case of RIs, the replicated region could also acquire a ΔLk that is induced by the inter-duplex crossings of the sister duplexes.

To understand the conformational properties of partially replicated DNA molecules in thermodynamic equilibrium, we used a mathematical approach, based on numerical simulations of RIs with the forks stalled at different sites before termination and without any mechanical impediment that avoids the propagation of torsional stress between both regions of the molecule through the replication forks (in a free state, after deproteinization). In each case, we generated a WLC ensemble of RIs with a certain degree of torsional strain accumulated either ahead or behind the forks, causing a non-vanishing torque. Replication forks were allowed to rotate until the torque balance was achieved, changing the distribution of the elastic energy between the replicated and un-replicated regions in order to find the thermodynamic equilibrium conformation using the Metropolis MC procedure. Our results confirmed a dynamic interchange of torsional stress between the two regions separated by the forks. The equilibrium distribution of torsional stress is related to the location of the replication forks relative to the termination sites.

Material and methods

The replication intermediate model

We modeled double-stranded DNA as a discrete worm-like chain (Frank-Kamenetskii et al., 1985; Vologodskii & Cozzarelli, 1993; Vologodskii et al., 1992). In this model DNA is represented by a series of rigid segments that are cylinders of constant length (3.4 nm) and thickness (3.0 nm), according to the value of the DNA effective diameter, impenetrable to each other. To model RIs using the WLC we created an un-

replicated region connected through replication forks to two circular molecules representing the sister duplexes of the replicated region. We modeled a circular plasmid DNA of 4440 bp where 25, 50 or 75% of the molecule has been replicated. In the case of the 50% replicated molecules, the un-replicated region as well as each nascent duplex in the replicated region were composed of 200 segments (2220 bp). In the case of 25% and 75% replicated molecules, the number of segments in each simulated curve was proportional to the length in bp of the simulated regions of the RI.

Figure 2 shows a half replicated molecule with DNA strands of the un-replicated and replicated parts modeled as a single and double coupled WLC, respectively. The un-replicated region (in blue) is connected through replication forks to two circular molecules representing the sister duplexes of the replicated region (in red). Newly replicated sister DNA duplexes could be multiply interlinked with (-) pre-catenanes and un-replicated DNA is (-) supercoiled. The topology of the RIs was modified by winding/unwinding the plectonemic structure of the un-replicated region or by braiding/unbraiding both nascent curves in the replicated portion. The number of supercoil crossings in the un-replicated region and the number of pre-catenane crossings in the replicated one determines the ΔLk of the RI. The mechanism to model replication forks rotation is shown in the inset of Figure 2. Red and blue dots represent the junction between consecutive segments while black dots correspond to the midpoints of the central axis between the red curves. Fork swiveling was achieved through rotation of both red curves around each other. In geometric terms, this was accomplished by applying the rotation of the two opposite points r'_i and r''_i (in each red curve) with respect to their midpoint MP_i . The segment $r'_i - MP_i - r''_i$ was rotated over an angle ϕ from its original orientation maintaining the point MP_i fixed. Depending on the direction, the rotation resulted in braiding or unbraiding of newly replicated curves, changing the topology of the replicated region. Through this mechanism, both forks were able to rotate at every step of the simulation procedure.

Energy fundamentals

The elastic potential energy model considered for simulated DNA molecules included the bending elastic energy, denoted by E_b , and the torsional energy, denoted by E_t (Vologodskii et al., 1992). Considering the WLC model, the bending energy is given by

$$E_b = k_B T \alpha \sum_{i=1}^N \theta_{i,i+1}^2, \quad (1)$$

where N denotes the number of segments, $\theta_{i,i+1}$ is the angular displacement between two consecutive segments, i and $i+1$, k_B is the Boltzmann constant, T is the absolute temperature (297 K), and α is the bending rigidity constant (Frank-Kamenetskii et al., 1985; Vologodskii & Cozzarelli, 1993; Vologodskii et al., 1992).

The torsional energy depends on the displacement of chain twist from its equilibrium value, ΔTw , and in this model it can be expressed by the equation $\Delta Tw = \Delta Lk - Wr$ (de Vries, 2005; Fuller, 1971), where ΔLk is the linking number difference of the simulated DNA. The value of ΔLk is considered as a constant parameter at each iteration of the simulation procedure (Vologodskii et al., 1992) and the Wr is a function of a single space curve that is necessary to compute after each trial move of the Metropolis MC method. Hence, the torsional energy E_t is defined as

$$E_t = \frac{2\pi^2 C}{L} \Delta Tw^2 = \frac{2\pi^2 C}{L} (\Delta Lk - Wr)^2, \quad (2)$$

where C is the torsion rigidity constant of the molecule ($4.50 \cdot 10^{-19}$ J.bp) and L is the DNA length expressed in bp (Vologodskii & Cozzarelli, 1993). Super-helical stress was relaxed in the nascent chains due to the single-stranded breaks. For this reason, the only component of the total energy of the replicated curves was the bending energy E_b .

Simulation procedure and computational performance

Using the model described above we created three curves, one that represents the un-replicated region and the other two corresponding to the nascent duplexes of the replicated region. The length of every curve was related to the percentage of replication that was simulated (25, 50 or 75%). The Metropolis MC procedure was applied to find the minimum elastic energy conformation and trial perturbations consisted in the rotation of an arbitrary number of adjacent segments of the WLC ensemble by a random angle around the straight line connecting two randomly chosen fixed points (Frank-Kamenetskii et al., 1985; Vologodskii & Cozzarelli, 1993; Vologodskii et al., 1992). The total number of turns of the replication forks, $Total_{Turns}$, and the total number of steps of the simulation, $Total_{Steps}$, determined the value of the rotation angle of the replication forks at every step, $\sigma = 2\pi \frac{Total_{Turns}}{Total_{Steps}}$. Hence, the number of turns at every step of the simulation procedure is given by $\sigma/2\pi$.

For every simulation, a set of initial conformations was created, each with a specific value of ΔLk distributed between the un-replicated and replicated parts. The basic

step of the simulation procedure for the redistribution of the ΔLk between the two regions of the RI was accomplished by changing the supercoiling of the un-replicated region in a value $\sigma/2\pi$ and rotating the replicated curves around each other with the angle σ . The algorithm was set so that the braiding of the replicated curves caused a loss of supercoils in the un-replicated region, and vice-versa, in order to keep a constant value of ΔLk . Finally, the Metropolis MC method was applied to the entire molecule in order to find the minimum elastic energy conformation in which the net torque is vanished. This basic step was repeated to reach a selected complexity or a specific energetic property.

Results

Simulation of the interchange of supercoiling and pre-catenation in replicating DNA

To study the interchange of torsional strain between both regions in half replicated DNA molecules in deproteinized conditions *in vitro*, we performed numerical simulations using the WLC model and the Metropolis MC method to analyze the elastic potential energy of 50% replicated RIs by migrating all torsional stress from one region to the other. First, we modeled RIs with 10 (-) supercoil crossings and devoid of pre-catenane crossings. The parental strands in the un-replicated region were topologically constrained due to supercoiling, whereas in the replicated region the sister duplexes were relaxed because swiveling of the free ends of newly synthesized chains allows the release of any super-helical stress. This is how half-replicated molecules would look *in vivo* if fork rotation was prevented except at termination (Schalbetter et al., 2015; Sundin & Varshavsky, 1980, 1981). Once all proteins are removed *in vitro*, the local torsional stress accumulated could migrate from the un-replicated to the replicated region and vice-versa by swiveling of the replication forks. In this way the supercoiling of the un-replicated region could lead to intertwining of the daughter DNA duplexes. The forks were rotated to allow the migration of all the (-) crossings from the un-replicated region to the replicated one. Figure 3A shows the energy shift after each rotation as a function of the crossing numbers of the replicated and un-replicated regions. It is clearly noted that the value of elastic energy of DNA supercoiling in the un-replicated region (in blue) decreased as the pre-catenation number of the replicated DNA increased (in red). The average elastic potential energy (calculated from 10 independent simulations) of the initial conformation was 34.2 Kcal/mol, 29.6 Kcal/mol for the un-replicated region and 4.7 Kcal/mol for the replicated portion. At the end, the elastic energy of the un-replicated region was reduced to 4.2 Kcal/mol while the elastic energy of the replicated region increased to 46.6 Kcal/mol. An example of the simulations performed is shown in Supplementary Video 1.

We also modeled half-replicated RIs with 20 (-) crossings in the replicated region and devoid of crossings in the un-replicated region. In this case, swiveling of replication fork induced the migration of all torsional stress to the un-replicated region. The initial values of elastic energy, resulting

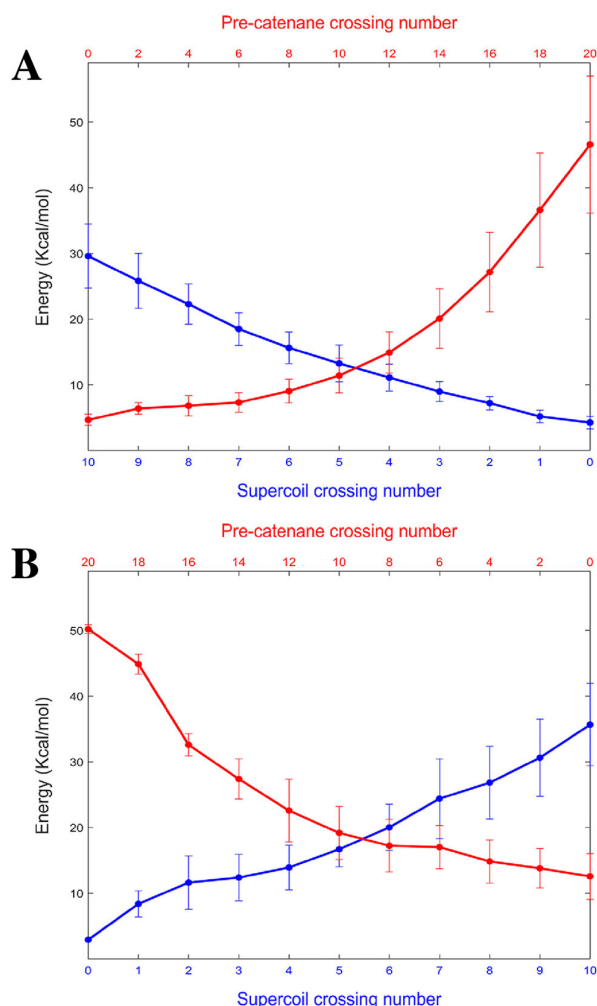


Figure 3. Fork rotation and migration of negative torsional stress between un-replicated and replicated regions. The topology of the RIs was modified through fork rotation either by changing the supercoiling of the un-replicated region or by rotating both nascent curves around each other in the replicated portion. Elastic energy was calculated after each rotation. (A): Profile of elastic energy corresponding to 50% replicated RIs with 10 (-) supercoil crossings in the un-replicated region and devoid of pre-catenanes. Forks were rotated to allow migration of all the (-) crossings from the un-replicated region to the replicated region. (B): Half replicated molecules with 20 (-) crossings in the replicated region and devoid of crossings in the un-replicated portion. Forks were rotated to allow the migration of all the (-) crossings from the replicated to the un-replicated region. Average values of elastic energy for the un-replicated (blue lines) and replicated (red lines) regions correspond to 10 independent simulations.

from 10 independent simulations, were 2.9 and 50.2 Kcal/mol for the un-replicated and replicated regions, respectively. The forks were rotated to allow the migration of all the (-) crossings from the replicated region to the un-replicated one. Figure 3B shows the energy shift after each rotation as a function of the crossing numbers of the replicated and un-replicated regions. It is clearly noted that here the value of elastic energy of pre-catenanes in the replicated region (in red) decreased as supercoiling in the un-replicated region increased (in blue). At the end, the elastic energy of the replicated region was reduced to 12.6 Kcal/mol while the elastic energy of the replicated region increased to 35.6 Kcal/mol. These results confirmed that our model for a RI *in vitro*

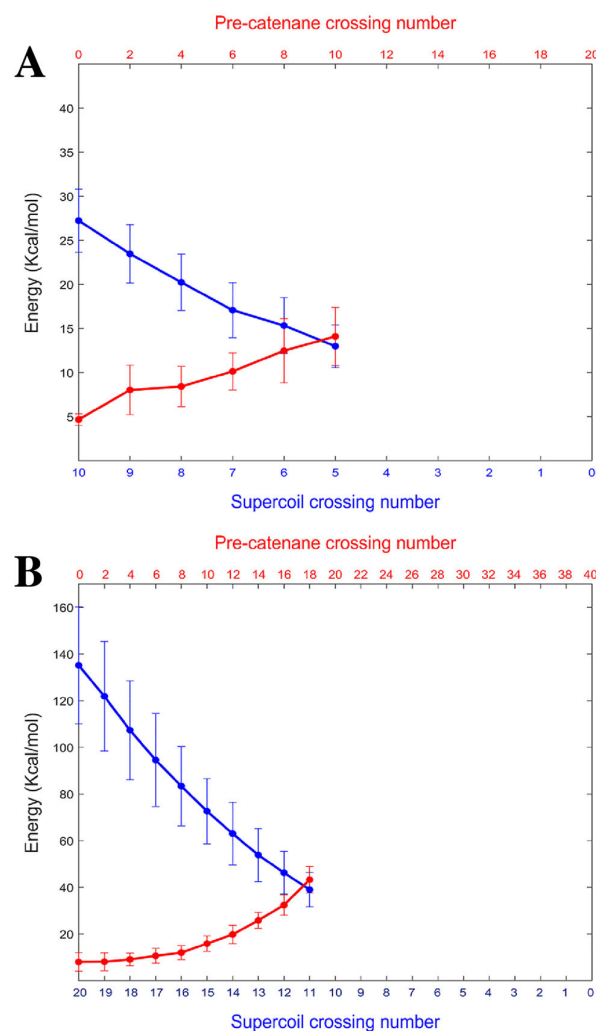


Figure 4. Thermodynamic equilibrium of 50% replicated negatively supercoiled DNA molecules. Swiveling of the replication fork allowed migration of torsional stress until both regions reached thermodynamic equilibrium. (A): Distribution of elastic energy in 50% replicated RIs with 10 (-) supercoil crossings in the un-replicated region and devoid of pre-catenanes. (B): Half-replicated molecules with 20 (-) supercoil crossings in the un-replicated region and devoid of pre-catenanes. Average values of elastic energy, corresponding to 10 independent simulations. Un-replicated (blue lines) and replicated (red lines) regions are represented as a function of the supercoil and pre-catenane crossing numbers.

was able to reproduce the migration of torsional strain from the un-replicated to the replicated region and vice-versa. An example of the simulations performed is shown in Supplementary Video 2.

Equilibrium conformation of half-replicated DNA molecules

In order to study the thermodynamic and conformational properties of the RIs in deproteinized conditions, we considered an ensemble of WLC-generated 50% replicated RIs with torsional strain accumulated either in the replicated or un-replicated region. In this case, swiveling of the replication forks allowed the migration of torsional stress until both regions reached thermodynamic equilibrium, i.e. the conformation in which the net torque of the molecule is vanished so the RI is in

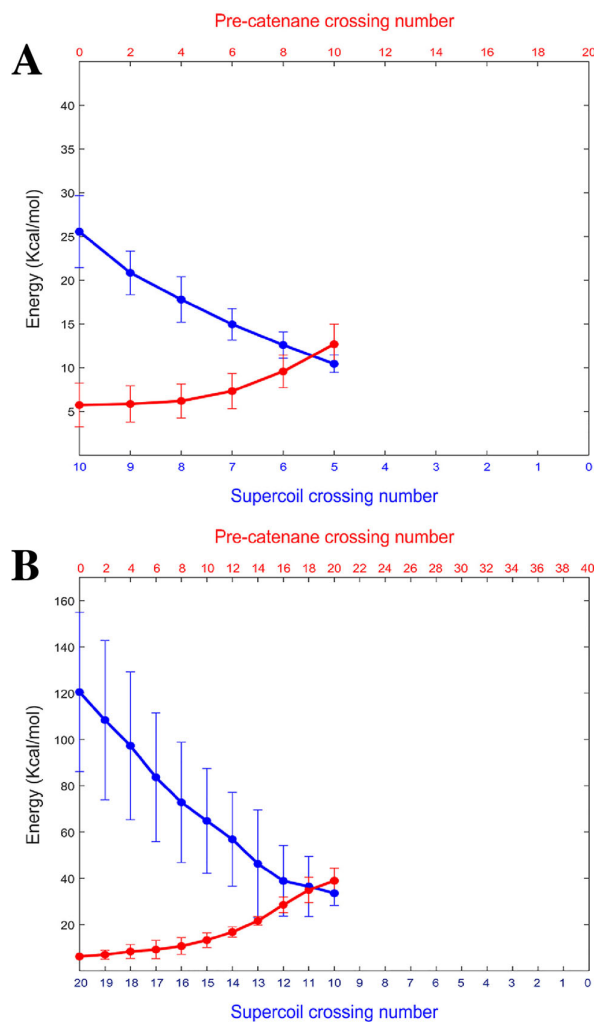


Figure 5. Thermodynamic equilibrium of 50% replicated positively supercoiled DNA molecules. Braiding of the replicated chains allowed the partial relaxation of (+) supercoils in the un-replicated region to achieve torque balance. (A): Distribution of elastic energy in 50% replicated RIs with 10 (+) supercoil crossings in the un-replicated region and devoid of pre-catenanes. (B): Half-replicated molecules with 20 (+) supercoil crossings in the un-replicated region and devoid of pre-catenanes. Average values of elastic energy, corresponding to 10 independent simulations, for the un-replicated (blue lines) and replicated (red lines) regions are represented as a function of the supercoil and pre-catenane crossing numbers.

mechanical equilibria. We modeled RIs with 10 (-) supercoils in the un-replicated region and devoid of pre-catenanes. Swiveling of the replication forks allowed partial release of the elastic energy accumulated in the un-replicated region to reach thermodynamic equilibrium ([Supplementary Video 3](#)). The elastic energy of both regions as a function of crossing numbers is shown in [Figure 4A](#). In the initial conformation the values of energy were 27.2 and 4.7 Kcal/mol for the un-replicated and replicated regions, respectively. At thermodynamic equilibrium, the final value of energy was 13.7 Kcal/mol for each region. Under this condition, partially replicated molecules with supercoiling in the un-replicated region and devoid of pre-catenanes reached thermodynamic equilibrium by migrating half of the crossings of the un-replicated region to form the same number of pre-catenanes. The equilibrium conformation, resulting from 10 independent simulations,

corresponded to a RI with 5.3 (-) crossings in the un-replicated region and 4.7 (-) pre-catenanes.

To expand this analysis, we repeated the simulation procedure with half replicated RIs showing 20 (-) supercoils in the un-replicated region and devoid of pre-catenanes. As in the previous case, swiveling of the replication fork allowed a partial release of the torsional stress stored in the un-replicated region to reach thermodynamic equilibrium ([Supplementary Video 4](#)). The elastic energy of both regions as a function of crossing numbers is shown in [Figure 4B](#). In the initial conformation the values of energy were 135.2 and 8.0 Kcal/mol for the un-replicated and replicated regions, respectively. At thermodynamic equilibrium the final value of potential elastic energy was 40.7 Kcal/mol for each region. Under this condition, partially replicated molecules with supercoiling in the un-replicated region and devoid of pre-catenanes reached thermodynamic equilibrium by migrating almost half of the crossings of the un-replicated region to form approximately the same number of pre-catenanes. The conformation with null net torque of the replication forks, resulting from 10 independent simulations, corresponded to a RI with 11.2 (-) crossings in the un-replicated region and 8.8 (-) pre-catenanes.

As previously mentioned, during DNA replication *in vivo*, unwinding of the parental DNA double-helix by DNA helicases causes a compensatory overwinding of the duplex ahead of the replication fork and this overwinding leads to positive (+) supercoiling (Kornberg & Baker, 1992). In 1980, Champoux and Been (Champoux & Been, 1980) proposed that in addition to the action of DNA topoisomerases in the un-replicated region, swiveling of the forks could reduce the (+) torsional tension accumulated immediately ahead of the replication fork at the expense of generating (+) intertwined pre-catenanes in the replicated region (Champoux & Been, 1980; Ullsperger et al., 1995). Therefore, to complete our analysis next we modeled RIs with 10 (+) supercoil crossings in the un-replicated region and devoid of pre-catenane crossings. Here, the parental strands in the un-replicated region were topologically constrained due to (+) supercoiling. Swiveling of the replication forks allowed partial release of the torsional stress accumulated in the un-replicated region to reach thermodynamic equilibrium ([Supplementary Video 5](#)). The elastic energy of both regions as a function of crossing numbers is shown in [Figure 5A](#). In the initial conformation the values of energy were 25.6 and 5.7 Kcal/mol for the un-replicated and replicated regions, respectively. At thermodynamic equilibrium, the final value of energy was 11.4 Kcal/mol for each region. Under this condition, partially replicated molecules with (+) supercoiling in the un-replicated region and devoid of pre-catenanes reached thermodynamic equilibrium by migrating half of the crossings of the un-replicated region to form the same number of (+) pre-catenanes. The equilibrium conformation, resulting from 10 independent simulations, corresponded to a RI with 5.4 (+) crossings in the un-replicated region and 4.6 (+) pre-catenanes.

We repeated the simulation procedure with half replicated RIs showing 20 (+) supercoils in the un-replicated region and devoid of pre-catenanes. As in the previous case, swiveling of the replication fork allowed a partial release of the torsional

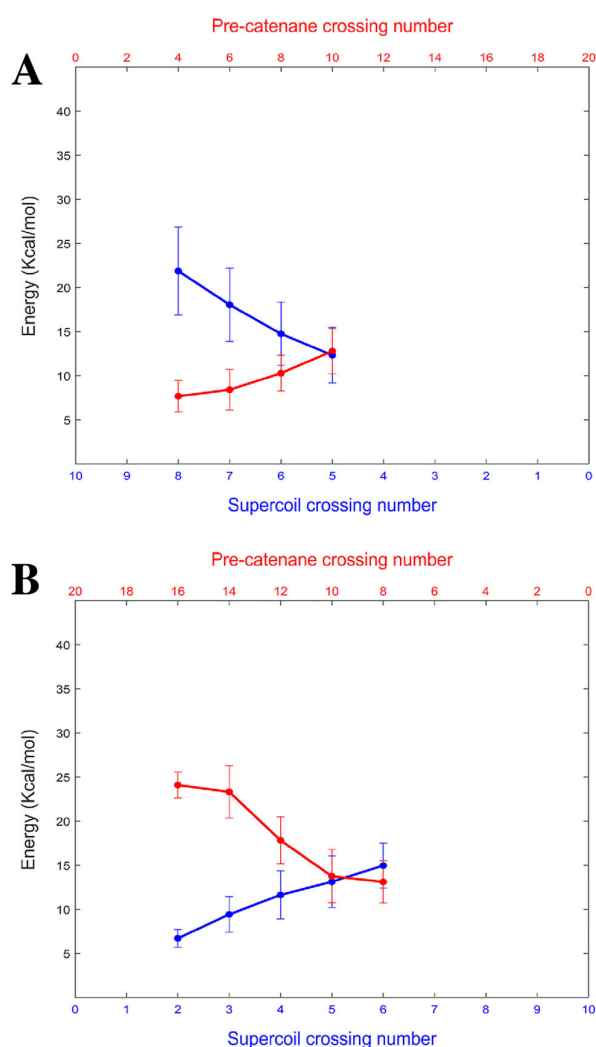


Figure 6. Distribution of torsional stress in 50% replicated RIs with different levels of topological complexity in the replicated and un-replicated regions. (A): Profiles of elastic energy corresponding to RIs with 8 (-) supercoil crossings in the un-replicated region and 4 (-) pre-catenane crossings in the replicated region. (B): RIs with 2 (-) supercoil crossings in the un-replicated region and 16 (-) pre-catenane crossings in the replicated portion. The replication forks were allowed to rotate until both regions reached thermodynamic equilibrium. Average values of elastic energy for the un-replicated (blue lines) and replicated (red lines) regions correspond to 10 independent simulations.

strain stored in the un-replicated region to reach thermodynamic equilibrium (Supplementary Video 6). The elastic energy of both regions as a function of crossing numbers is shown in Figure 5B. In the initial conformation the values of energy were 120.5 and 6.2 Kcal/mol for the un-replicated and replicated regions, respectively. At thermodynamic equilibrium the final value of energy was 35.8 Kcal/mol for each region. Under this condition, partially replicated molecules with (+) supercoiling in the un-replicated region and devoid of pre-catenanes reached thermodynamic equilibrium by migrating almost half of the crossings of the un-replicated region to form approximately the same number of (+) pre-catenanes. The equilibrium conformation, resulting from 10 independent simulations, corresponded to a RI with 10.8 (+) crossings in the un-replicated region and 9.2 (+) pre-catenanes.

Next we extended our analyses to show the distribution of mechanical stress in half-replicated RIs with different number of crossings in the un-replicated and replicated regions. In the first case, the initial conformation was a RI with 8 (-) supercoils in the un-replicated region and 2 (-) pre-catenanes in the replicated region. Rotation of the replication forks allowed the migration of torsional stress to balance the torque between both regions (Supplementary Video 7). The elastic energy as a function of the crossing numbers in both regions is shown in Figure 6A. There was a gradual loss of crossings in the un-replicated region with a synchronized acquisition of new intertwining between the two daughter duplexes. The initial values of elastic energy for the un-replicated and replicated regions were 21.9 and 7.7 Kcal/mol, respectively. Fork rotation stopped when both regions achieved the same elastic energy (12.6 Kcal/mol). The equilibrium conformation corresponded to a RI with 5.1 (-) supercoils and 4.9 (-) pre-catenanes. In the opposite situation, an ensemble of RIs with 2 (-) crossings in the un-replicated region and 16 (-) crossings in the replicated region, redistribution of torsional stress resulted in an equilibrium conformation with 5.3 (-) supercoils and 4.7 (-) pre-catenanes (Figure 6B). The initial values of energy of the un-replicated and replicated regions were 6.7 and 24.1 Kcal/mol, respectively. Total elastic potential energy at the equilibrium conformation was 27.2 Kcal/mol, 13.6 Kcal/mol for each region (Supplementary Video 8). These results confirmed that, for all the initial conformations tested, rotation of the replication forks to reach thermodynamic equilibrium allows a homogeneous distribution of torsional stress in 50% replicated molecules.

Distribution of the elastic energy of RIs with the forks stalled at different sites before termination

We further tested our model by probing an ensemble of RIs with different relative sizes of replicated and un-replicated regions. Inspired by *in vivo* experiments (Cebrian et al., 2015), we analyzed the distribution of elastic energy in a 25% replicated RI with 10 (-) supercoils in the un-replicated region and devoid of pre-catenanes. Fork rotation allowed the migration of torsional stress to reach thermodynamic equilibrium. The energy shift after each rotation of the forks is shown in Figure 7A. In the initial conformation, the elastic energy of the replicated region was 6.1 Kcal/mol and it was increased to 12.5 Kcal/mol at the equilibrium conformation. In the un-replicated region, there was a coordinated decline of energy from 21.1 to 12.5 Kcal/mol. The conformation in which the net torque is vanished, resulting from 10 independent simulations, corresponded to a RI with 6.8 (-) crossings in the un-replicated region and 3.2 (-) pre-catenanes in the replicated part. An example of the redistribution of torsional stress is shown in Supplementary Video 9.

Simulation was also used to analyze 75% replicated RIs with 10 (-) supercoils in the un-replicated region and devoid of crossings in the replicated region. Distribution of torsional stress by rotation of the replication forks to achieve torque balance occurred in such a way that most of the (-)

supercoils switched to (-) pre-catenanes. A profile of the elastic energy in both regions is shown in Figure 7B. Fork rotation stopped when both regions achieved thermodynamic equilibrium (15.6 Kcal/mol). The equilibrium conformation, resulting from 10 independent simulations, corresponded to a RI with 3.5 (-) crossings in the un-replicated region and 6.5 (-) pre-catenanes in the replicated region (Supplementary Video 10). These results demonstrated that the equilibrium distribution of torsional stress between both regions is related to the ratio between the sizes of both regions.

Figure 8 summarizes the initial and final conformations of some simulated cases (see also Supplementary Table 1). Altogether, these results showed that the equilibrium conformation is dynamic as it changes according to relative sizes of the replicated and un-replicated portions of the RI. For all the cases studied, the initial energy observed was greater than the final one when RIs reached thermodynamic equilibrium.

Discussion

Once replication terminates the newly synthesized sister duplexes appear heavily intertwined (Liu et al., 2009; Martinez-Robles et al., 2009; Schalbetter et al., 2015; Sundin & Varshavsky, 1980, 1981). Although DNA rings may be catenated *in vitro* (Baldi et al., 1980; Kreuzer & Cozzarelli, 1980), it was clearly shown that *in vivo* catenanes derive from pre-catenanes formed during replication (Adams et al., 1992; Martinez-Robles et al., 2009; Peter et al., 1998; Ullsperger et al., 1995). In turn, these pre-catenanes are formed due to the diffusion of torsional stress from the un-replicated to the replicated region of partially replicated molecules driven by fork swiveling (Cebrian et al., 2015; Champoux & Been, 1980; Ullsperger et al., 1995). By making direct torque measurements it was found that in naked DNA (without nucleosomes), torsional stress generated by replication is partitioned primarily behind the replisome. In contrast, in eukaryotes, a single chromatin fiber (like the un-replicated region) is torsionally soft and a braided fiber (like the replicated chains) is torsionally stiff, meaning that the superhelical stress on chromatin is preferentially stored in the un-replicated region (Le et al., 2019). However, when and how do replication forks rotate to allow the passage of torsional stress from the un-replicated to the replicated region is still poorly understood.

A very important aspect to be considered is the topological sign of supercoiling in the un-replicated region and pre-catenation in the replicated one. In all the cases studied so far *in vivo*, catenane's crossings in fully replicated sister duplexes are (+) (Sundin & Varshavsky, 1980, 1981). As these catenanes derive from pre-catenanes, their crossings should be (+) as well. We previously discussed that *in vivo*, pre-catenation forms due to the diffusion of torsional stress from the un-replicated to the replicated regions of partially replicated molecules facilitated by fork swiveling (Cebrian et al., 2015; Champoux & Been, 1980; Ullsperger et al., 1995). Therefore, we must conclude that the (+) pre-catenanes formed *in vivo* derive from (+) supercoiling that transiently

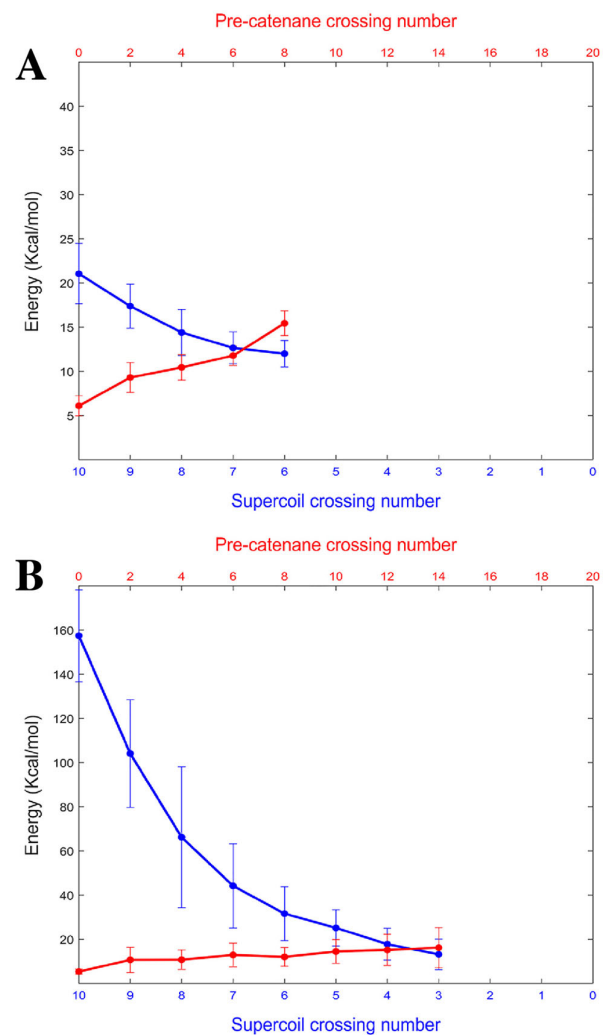


Figure 7. Distribution of torsional stress in 25 and 75% replicated negatively supercoiled DNA molecules. (A): Profiles of elastic energy as a function of the crossing numbers corresponding to 25% replicated RI, with 10 (-) supercoils in the un-replicated region and devoid of pre-catenanes. (B): 75% replicated RI, with 10 (-) supercoil crossings in the un-replicated region and devoid of pre-catenanes. Average values of elastic energy for the un-replicated (blue lines) and replicated (red lines) correspond to 10 independent simulations.

accumulate immediately ahead of progressing forks. In prokaryotes, topoisomerase IV (Topo IV) is ~ 20 times more efficient eliminating (+) than (-) crossings (Crisona et al., 2000; Stone et al., 2003). It was also shown that bacterial plasmids are more torsionally tensioned in the absence of Topo IV (Cebrian et al., 2015). These observations indicate that in bacterial plasmids interconversion of (+) supercoiling from the un-replicated region and pre-catenane's (+) crossings occurs continuously during replication. On the other hand, in the budding yeast, Schalbetter and co-workers showed that fork rotation and the formation of pre-catenanes during replication *in vivo* are actively inhibited by Timeless/Tof1 and Tipin/Csm3 and only occur in hard-to-replicate contexts such as stable protein-DNA fragile sites and termination (Schalbetter et al., 2015). In any case, during DNA isolation once all proteins are removed the replication forks of partially replicated molecules can rotate freely and the only force driving this

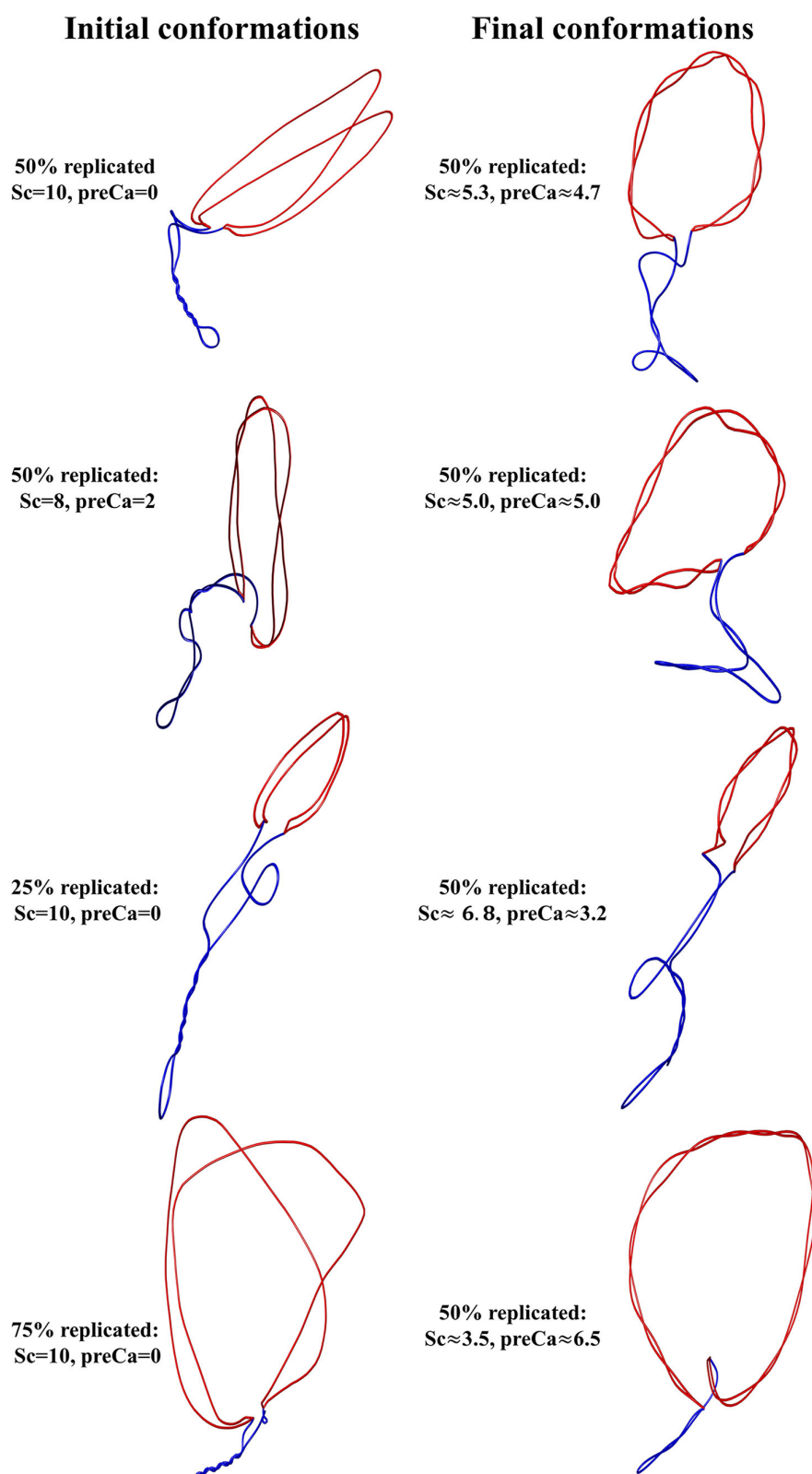


Figure 8. The initial and final conformations in several of the simulations tested. For a 50% replicated RI starting with 10 (-) supercoil crossings in the un-replicated region and devoid of pre-catenanes. For a 50% replicated RI starting with 8 (-) supercoil crossings in the un-replicated region and 4 (-) pre-catenane crossings in the replicated region. For a 25% replicated RI starting with 10 (-) supercoil crossings in the un-replicated region and devoid of pre-catenanes. And for a 75% replicated RI starting with 10 (-) supercoil crossings in the un-replicated region and devoid of pre-catenanes. In all cases the simulation procedure was stopped once thermodynamic equilibrium was achieved.

swiveling is the energy accumulated by torsional stress in the un-replicated and replicated regions. Taking this into account, here we developed a worm-like chain model to study the interchange of supercoiling and pre-catenation in partially replicated molecules in a deproteinized condition. Our study showed that, at thermodynamic equilibrium, torsional stress distribution depends on the progression of the replication fork, as seen for 25, 50 and 75% partially replicated molecules. We observed that the torsional stress accumulated in the un-replicated and replicated regions depends on their relative sizes, i.e. the extent of replication. These results confirmed that the equilibrium distribution of ΔLk between replicated and un-replicated regions should approximate to the ratio of their length in base pairs (Ullsperger et al., 1995). If the forks were only allowed to rotate as termination approaches most of the replicon would have already been replicated. On the contrary, if forks are allowed to swivel continuously, interchange of supercoils and pre-catenanes would continuously change as replication advances. This would certainly add another twist to the dynamics of DNA replication.

To our knowledge, this is the first study of the interchange of torsional strain between the un-replicated and replicated regions in replicating DNA using a numerical approach. In our numerical simulations, the energy gradient between both regions was significantly high at the beginning, when all the torsional stress was accumulated in one region (replicated or un-replicated). Once rotation of the replication fork was performed to achieve torque balance, the elastic energy difference between both regions was reduced as the torsional stress migrated from the region with higher complexity to the more relaxed one. The rate of change of the energy gradient between both regions of the RI was faster at the beginning and gradually slowed down as the RI approached to thermodynamic equilibrium. The imbalance in torque, as a consequence of the imposed local torsional strain, could be considered as the driving force of fork's rotation as it regulates the speed of the process. Here we propose that the distribution of torsional stress between the un-replicated and replicated regions drives fork swiveling allowing the interchange of supercoiling and pre-catenation during DNA replication.

Our model mimics partially replicated DNA strands with replication forks stalled at different sites before termination in deproteinized conditions. DNA strands of replicated and un-replicated regions were modeled as a single and double coupled WLC, respectively, separated by swiveling forks. To study the dynamic of the distribution of torsional strain between the two regions separated by the forks, we found the thermodynamic equilibrium conformation using Metropolis MC. The data obtained revealed that the equilibrium conformational properties of the simulated RIs were independent of the initial conditions or the topological sign of supercoiling (Figures 4–6).

The model proposed here assumes that in order to study the exchange between supercoiling and pre-catenation, we need to consider bending and torsional components of the elastic energy. DNA subjected to bending energy is often described by WLC modeling. This bending energy is faithfully

taken into by equation (1). On the other hand, simple WLC do not explicitly include twist degrees freedom. In the worm-like chain model, the elementary rods that form the simulated molecule of DNA usually have two degrees of freedom; bending and twisting. Twisting degree of freedom is associated with the rotation of one rod of the WLC respect to the next one. Bending degree of freedom is related to the angle between the axis of one rod respect to the axis of the adjacent rod (de Vries, 2005; Vologodskii et al., 1992). There are different approaches to calculate the torsional component of energy, including the twistable worm-like chain (TWLC) (Nomidis et al., 2019) or helical worm-like chain (Liu et al., 2011). Since we are not interested in the local twisting dynamics, or in the un-homogeneities of the distribution of the twist along the chain, the integration of the local twist angles along the chain can be used to define the torsional component of the elastic energy (de Vries, 2005). Here we used the expression of twist energy proposed by Vologodskii and co-workers (Vologodskii et al., 1992) that, despite its simplicity, describes many thermodynamic properties of DNA. This implementation of the WLC model has several assumptions and approximations, i.e. the approximation of the DNA as an isotropic elastic chain, so local sequence-dependent variations in structure are ignored; no consideration of supercoiling-driven structural transitions such as the introduction of cruciforms and Z-DNA; assumption that torsional and bending rigidity are independent; approximation of electrostatic interactions between DNA segments by cylinders whose effective diameter accounts for excluded volume effects (Vologodskii et al., 1992). Despite of these limitations, the simulation procedure used here reproduces physical properties of supercoiled and catenated DNA (Martinez-Robles et al., 2009; Vologodskii & Cozzarelli, 1993; Vologodskii et al., 1992), referring specifically to the conformational properties of the un-replicated and replicated regions.

The Metropolis Monte Carlo method implemented in this study finds the potential elastic energy for each portion of the RI, which depends on the distribution of the elastic deformation between both regions. In all cases where the simulation procedure was stopped when both regions reached thermodynamic equilibrium there was a difference in total elastic energy between the initial and final conformations because the energy profiles presented do not include other components of total energy, like the kinetic energy or the dissipated energy. Such difference was more evident in molecules accumulating more torsional stress, like the one shown in Figure 4B ($\Delta Lk = -20$). In the case of the 75% replicated molecule (Figure 7B & Supplementary Video 10), the size of the un-replicated region is significantly shorter and the bending effect is greater. As a consequence, the initial energy was higher than that for 50% replicated molecules, which in turn was higher than the initial energy of 25% replicated RIs. Independently of the relative size of both regions, the final conformations had similar elastic energy values. In the case where the final energy was greater than the initial one (Figure 3A & Supplementary Video 1) the difference observed was due to the simulation procedure resulting in a final conformation that could not have been

reached through a spontaneous process. Nonetheless, it was of interest to study the energy properties at this state, too.

Another theoretical approximation for numerical simulations of replication intermediates could be tested in future works with a model that includes more degrees of freedom than those exposed in the model implemented here, i.e. models in which local variations of the torsion angles are taken into account. Alternatively, as a future task, computational methods such as molecular dynamics (MD) simulations could be used to predict replication fork dynamics and the evolution of the simulated molecules through the phase space.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was sustained by grant PINV15-573 from the Paraguayan CONACYT-PROCIENCIA program to MJFN and BFU2014-56835 from the Spanish Ministerio de Economía y Competitividad to JBS. The authors acknowledge the critics and continuous support of Prof. Andrzej Stasiak.

References

- Adams, D. E., Shekhtman, E. M., Zechiedrich, E. L., Schmid, M. B., & Cozzarelli, N. R. (1992). The role of topoisomerase-IV in partitioning bacterial replicons and the structure of catenated intermediates in DNA replication. *Cell*, 71(2), 277–288. doi:10.1016/0092-8674(92)90356-H
- Baldi, M. I., Benedetti, P., Mattoccia, E., & Tocchini-Valentini, G. P. (1980). In vitro catenation and decatenation of DNA and a novel eucaryotic ATP-dependent topoisomerase. *Cell*, 20(2), 461–467. doi:10.1016/0092-8674(80)90632-7
- Burnier, Y., Dorier, J., & Stasiak, A. (2008). DNA supercoiling inhibits DNA knotting. *Nucleic Acids Research*, 36(15), 4956–4963. doi:10.1093/nar/gkn467
- Cebrian, J., Castan, A., Martinez, V., Kadomatsu-Hermosa, M. J., Parra, C., Fernandez-Nestosa, M. J., Schaerer, C., Hernandez, P., Krimer, D. B., & Schvartzman, J. B. (2015). Direct evidence for the formation of precatenanes during DNA replication. *Journal of Biological Chemistry*, 290(22), 13725–13735. doi:10.1074/jbc.M115.642272
- Champoux, J. J., & Been, M. D. (1980). Topoisomerases and the swivel problem. In B. Alberts (Ed.), *Mechanistic studies of DNA replication and genetic recombination* (pp. 809–815). Academic Press.
- Crisona, N. J., Strick, T. R., Bensimon, D., Croquette, V., & Cozzarelli, N. R. (2000). Preferential relaxation of positively supercoiled DNA by E. coli topoisomerase IV in single-molecule and ensemble measurements. *Genes & Development*, 14(22), 2881–2892. doi:10.1101/gad.838900
- de Vries, R. (2005). Evaluating changes of writhe in computer simulations of supercoiled DNA. *Journal of Chemical Physics*, 122(6), 064905. doi:10.1063/1.1846052
- Frank-Kamenetskii, M. D., Lukashin, A. V., Anshelevich, V. V., & Vologodskii, A. V. (1985). Torsional and bending rigidity of the double helix from data on small DNA rings. *Journal of Biomolecular Structure and Dynamics*, 2(5), 1005–1012. doi:10.1080/07391102.1985.10507616
- Fuller, F. B. (1971). The writhing number of a space curve. *Proceedings of the National Academy of Sciences of Sciences*, 68(4), 815–819. doi:10.1073/pnas.68.4.815
- Funnell, B. E., Baker, T. A., & Kornberg, A. (1986). Complete enzymatic replication of plasmids containing the origin of the Escherichia coli chromosome. *The Journal of Biological Chemistry*, 261(12), 5616–5624.
- Klenin, K., & Langowski, J. (2000). Computation of writhe in modeling of supercoiled DNA. *Biopolymers*, 54(5), 307–317. doi:10.1002/1097-0282(20001015)54:5<307::AID-BIP20>3.0.CO;2-Y
- Kornberg, A., & Baker, T. A. (1992). *DNA replication* (2nd ed.). W.H. Freeman and Co.
- Kreuzer, K. N., & Cozzarelli, N. R. (1980). Formation and resolution of DNA catenanes by DNA gyrase. *Cell*, 20(1), 245–254. doi:10.1016/0092-8674(80)90252-4
- Le, T. T., Gao, X., Park, S. H., Lee, J., Inman, J. T., Lee, J. H., Killian, J. L., Badman, R. P., Berger, J. M., & Wang, M. D. (2019). Synergistic coordination of chromatin torsional mechanics and topoisomerase activity. *Cell*, 179(3), 619–631. doi:10.1016/j.cell.2019.09.034
- Liu, Z., Deibler, R. W., Chan, H. S., & Zechiedrich, L. (2009). The why and how of DNA unlinking. *Nucleic Acids Research*, 37(3), 661–671. doi:10.1093/nar/gkp041
- Liu, Y., Perez, T., Li, W., Gunton, J. D., & Green, A. (2011). Statistical mechanics of helical wormlike chain model. *Journal of Chemical Physics*, 134(6), 065107. doi:10.1063/1.3548885
- Marians, K. J., Minden, J. S., & Parada, C. (1986). Replication of superhelical DNAs in vitro. *Prog Nucleic Acids Res Mol Biol*, 33, 111–140.
- Mariezcurrera, A., & Uhlmann, F. (2017). Observation of DNA intertwining along authentic budding yeast chromosomes. *Genes & Development*, 31(21), 2151–2161. doi:10.1101/gad.305557.117
- Martinez-Robles, M. L., Witz, G., Hernandez, P., Schvartzman, J. B., Stasiak, A., & Krimer, D. B. (2009). Interplay of DNA supercoiling and catenation during the segregation of sister duplexes. *Nucleic Acids Research*, 37(15), 5126–5137. doi:10.1093/nar/gkp530
- Nomidis, S. K., Skoruppa, E., Carlon, E., & Marko, J. F. (2019). Twist-bend coupling and the statistical mechanics of the twistable wormlike-chain model of DNA: Perturbation theory and beyond. *Physical Review E*, 99(3-1), 032414doi:10.1103/PhysRevE.99.032414
- Peter, B. J., Ullsperger, C., Hiasa, H., Marians, K. J., & Cozzarelli, N. R. (1998). The structure of supercoiled intermediates in DNA replication. *Cell*, 94(6), 819–827. doi:10.1016/S0092-8674(00)81740-7
- Racko, D., Benedetti, F., Dorier, J., Burnier, Y., & Stasiak, A. (2017). Molecular dynamics simulation of supercoiled, knotted, and catenated DNA molecules, including modeling of action of DNA gyrase. *Methods in Molecular Biology*, 1624, 339–372. doi:10.1007/978-1-4939-7098-8_24
- Rawdon, E. J., Dorier, J., Racko, D., Millett, K. C., & Stasiak, A. (2016). How topoisomerase IV can efficiently unknot and decatenate negatively supercoiled DNA molecules without causing their torsional relaxation. *Nucleic Acids Research*, 44(10), 4528–4538. doi:10.1093/nar/gkw311
- Schalbetter, S. A., Mansoubi, S., Chambers, A. L., Downs, J. A., & Baxter, J. (2015). Fork rotation and DNA precatenation are restricted during DNA replication to prevent chromosomal instability. *Proceedings of the National Academy of Sciences*, 112(33), E4565–E4570. doi:10.1073/pnas.1505356112
- Schvartzman, J. B., Hernandez, P., Krimer, D. B., Dorier, J., & Stasiak, A. (2019). Closing the DNA replication cycle: From simple circular molecules to supercoiled and knotted DNA catenanes. *Nucleic Acids Research*, 47(14), 7182–7198. doi:10.1093/nar/gkz586
- Schvartzman, J. B., Martinez-Robles, M. L., Hernandez, P., & Krimer, D. B. (2013). The benefit of DNA supercoiling during replication. *Biochemical Society Transactions*, 41(2), 646–651. doi:10.1042/BST20120281
- Schvartzman, J. B., & Stasiak, A. (2004). A topological view of the replicon. *EMBO Reports*, 5(3), 256–261. doi:10.1038/sj.embor.7400101
- Stone, M. D., Bryant, Z., Crisona, N. J., Smith, S. B., Vologodskii, A., Bustamante, C., & Cozzarelli, N. R. (2003). Chirality sensing by Escherichia coli topoisomerase IV and the mechanism of type II topoisomerases. *Proceedings of the National Academy of Sciences of Sciences*, 100(15), 8654–8659. doi:10.1073/pnas.1133178100
- Sundin, O., & Varshavsky, A. (1980). Terminal stages of SV40 DNA replication proceed via multiply intertwined catenated dimers. *Cell*, 21(1), 103–114. doi:10.1016/0092-8674(80)90118-X
- Sundin, O., & Varshavsky, A. (1981). Arrest of segregation leads to accumulation of highly intertwined catenated dimers dissection of the final stages of SV40 DNA replication. *Cell*, 25(3), 659–669. doi:10.1016/0092-8674(81)90173-2
- Ullsperger, C., Vologodskii, A. A., & Cozzarelli, N. R. (1995). Unlinking of DNA by topoisomerases during DNA replication. In D. M. J. Lilley & F.

- Eckstein (Eds.), *Nucleic acids and molecular biology* (pp. 115–142). Springer-Verlag.
- Vologodskii, A. (1992). *Topology and physics of circular DNA*. CRC Press.
- Vologodskii, A. (2011). Unlinking of supercoiled DNA catenanes by type IIA topoisomerases. *Biophysical Journal*, 101(6), 1403–1411. doi:[10.1016/j.bpj.2011.08.011](https://doi.org/10.1016/j.bpj.2011.08.011)
- Vologodskii, A. V., & Cozzarelli, N. R. (1993). Monte Carlo analysis of the conformation of DNA catenanes. *Journal of Molecular Biology*, 232(4), 1130–1140. doi:[10.1006/jmbi.1993.1465](https://doi.org/10.1006/jmbi.1993.1465)
- Vologodskii, A. V., Levene, S. D., Klenin, K. V., Frank-Kamenetskii, M., & Cozzarelli, N. R. (1992). Conformational and thermodynamic properties of supercoiled DNA. *Journal of Molecular Biology*, 227(4), 1224–1243. doi:[10.1016/0022-2836\(92\)90533-P](https://doi.org/10.1016/0022-2836(92)90533-P)
- Vologodskii, A., & Rybenkov, V. V. (2009). Simulation of DNA catenanes. *Physical Chemistry Chemical Physics*, 11(45), 10543–10552. doi:[10.1039/b910812b](https://doi.org/10.1039/b910812b)
- Wasserman, S. A., White, J. H., & Cozzarelli, N. R. (1988). The helical repeat of double-stranded DNA varies as a function of catenation and supercoiling. *Nature*, 334(6181), 448–450. doi:[10.1038/334448a0](https://doi.org/10.1038/334448a0)
- Witz, G., & Stasiak, A. (2010). DNA supercoiling and its role in DNA decatenation and unknotting. *Nucleic Acids Research*, 38(7), 2119–2133. doi:[10.1093/nar/gkp1161](https://doi.org/10.1093/nar/gkp1161)

Topology of Partially Replicated DNA Molecules

Victor Martínez¹

Facultad Politécnica. UNA, San Lorenzo, Paraguay

Christian Schaerer²

Facultad Politécnica. UNA, San Lorenzo, Paraguay

María José Fernández-Nestosá³

Facultad Politécnica. UNA, San Lorenzo, Paraguay

Pablo Hernández⁴

Centro de Investigaciones Biológicas, CSIC, Madrid, España

Dora B. Krimer⁵

Centro de Investigaciones Biológicas, CSIC, Madrid, España

Jorge B. Schwartzman⁶

Centro de Investigaciones Biológicas, CSIC, Madrid, España

How the topology of DNA changes during replication is still poorly understood. Every DNA molecule in a partial replication state can be modeled as two separated topological domains, the unreplicated and the replicated region, physically connected by replication forks. The balance of energy between these regions has a key role during DNA replication. The topology of a DNA molecule during the replication process is regulated by enzymes called topoisomerases, which are molecular targets for a variety of drugs like antibiotics and anticancer drugs. The processivity of the topoisomerases depends on the conformational and energetic properties of the replicating DNA molecule. These properties of the replication intermediates are not well characterized.

The topological and energetic characterization of circular DNA plasmids depends on the Linking number (Lk), which is an integer that determines the total number of turns of the double helix. Lk is a topological invariant and, according to the Calugareanu-White-Fuller relation, $Lk = Wr + Tw$, where Tw (Twist) determines the degree of torsional deformation of the double helix and Wr (Writhe) is a number that describes the three-dimensional folding of the molecule. The topology of partially replicated plasmids also depends on the pre-catenation number that describes the interlinking number between the two sister chains and equals half of the signed sum of intermolecular nodes. In this work, we performed numerical simulations to qualitatively identify the variation of energy and

¹vmartizchamorro@gmail.com

²chris.schaerer@gmail.com

³mjfernandez@pol.una.py

⁴p.hernandez@cib.csic.es

⁵dbkrimer@cib.csic.es

⁶schwartzman@cib.csic.es

characterize the conformational equilibrium of partial replicated molecules (Replication Intermediates, RIs). To this end, partially replicated circular DNA molecules were considered. We use the wormlike chain model, which has been probed to give a very accurate approximation to experimental observations. The potential energy was calculated as the sum of two elastic potentials: the first one due to the bending of the molecule and the second one representing the twisting effects of the double helix around itself. The model used to simulate a RI was composed of an unreplicated region connected through replication forks to two circular molecules representing the nascent chains of the replicated region. Newly replicated sister DNA molecules can be multiply interlinked with left-handed (LH) precatenanes and unreplicated can be supercoiled region with right-handed (RH) plectonemes. The topology of the RI was modified through fork rotation either by winding or unwinding the plectonemic structure in the unreplicated region or by rotating both nascent curves around each other in the replicated region. To analyze the stability of the equilibrium conformation, the numerical simulation was performed using the Metropolis Monte Carlo method (MMC), starting from several initial topological conformations. At every MMC step we tested the topological invariance property imposed by the molecule linking number. When a local equilibrium conformation was obtained, the algorithm modified the supercoiling of the unreplicated region and performed equivalent rotations of the replication fork, inducing a modification in the precatenation of the nascent chains. As a result, a new local equilibrium conformation was obtained with MMC in order to evaluate its energy.

Our results confirmed that the balance of energy between the unreplicated and replicated regions drives fork swiveling, allowing the interchange and redistribution of supercoiling and precatenation during DNA replication. In addition, the final equilibrium state of energy was asymptotically stable for all the initial topological conformations tested, and it was robust in terms of energy perturbations introduced by the MMC method.

References

- [1] J. Cebrián, A. Castán, V. Martínez, M.J. Kadomatsu-Hermosa, C. Parra, M. J. Fernández-Nestosa, C. Schaerer, P. Hernández, D. B. Krimer and J. B. Schwartzman, Direct Evidence for the Formation of Precatenanes during DNA Replication, *J Biol Chem*, 290:13725-13735, 2015.
- [2] B. J. Peter, C. Ullsperger, H. Hiasa, K. J. Mariani, and N.R. Cozzarelli, The structure of supercoiled intermediates in DNA replication, *Cell*, 94:819-827, 1998.
- [3] J. B. Schwartzman, M. L. Martínez-Robles, P. Hernández, D. B. Krimer, The benefit of DNA supercoiling during replication, *Biochemical Society Transactions*, 41:646-651, 2013.
- [4] A. V. Vologodskii, and N.R. Cozzarelli, Monte Carlo analysis of the conformation of DNA catenanes. *Journal of molecular biology*, 232:1130-1140, 1993.