



UNIVERSITAT_{DE}
BARCELONA

Estudio topológico del ADN nucleosomal y su paradoja del número de enlace

Joana Segura Martinez



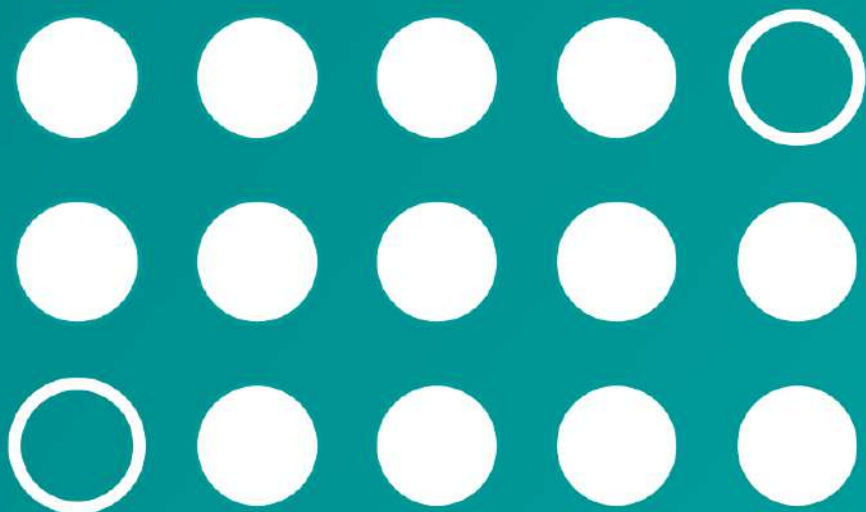
Aquesta tesi doctoral està subjecta a la llicència **Reconeixement 4.0. Espanya de Creative Commons.**

Esta tesis doctoral está sujeta a la licencia **Reconocimiento 4.0. España de Creative Commons.**

This doctoral thesis is licensed under the **Creative Commons Attribution 4.0. Spain License.**

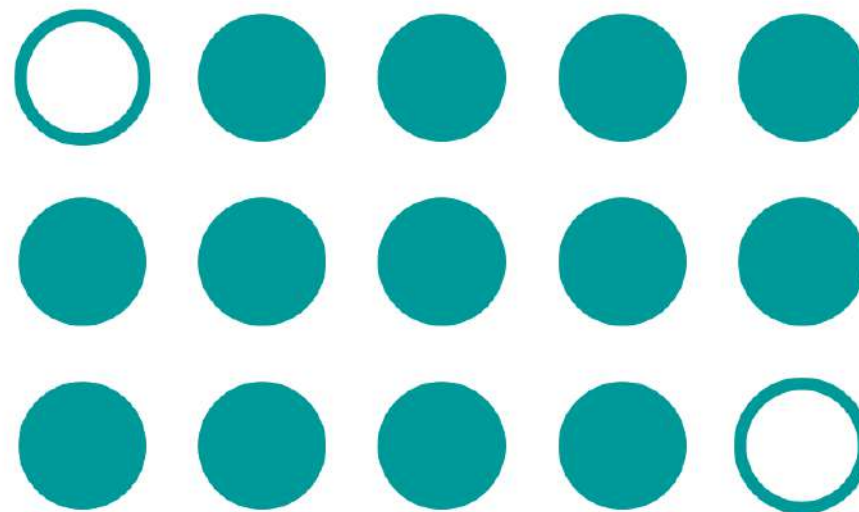
En los años 80 del siglo pasado, diversos estudios experimentales establecieron que la unidad estructural de la cromatina, el nucleosoma, estabilizaba una diferencia del número de enlace (ΔLk) igual a -1.0 /por nucleosoma. Este valor, sin embargo, difería con el valor teórico estimado mediante los estudios estructurales de los nucleosomas, que concluían un valor ΔLk cercano a -2 . Esta incongruencia generó la llamada "Paradoja del número de enlace del ADN nucleosomal", que ha estado objeto de debate hasta la actualidad.

En esta tesis, se presenta un minucioso estudio y revisión de la topología del nucleosoma aportando datos experimentales que esclarecen su paradoja.



TESIS DOCTORAL 2020

ESTUDIO TOPOLÓGICO DEL ADN NUCLEOSOMAL Y SU PARADOJA DEL NÚMERO DE ENLACE



Joana Segura Martinez



UNIVERSITAT DE
BARCELONA

TESIS DOCTORAL

ESTUDIO TOPOLÓGICO DEL ADN NUCLEOSOMAL Y SU PARADOJA DEL NÚMERO DE ENLACE

Memoria presentada por Joana Segura Martinez para optar al grado de

Doctora por la Universidad de Barcelona

Programa de Genética

Este trabajo ha sido realizado bajo la dirección del Dr. Joaquim Roca Bosch

Unidad de Biología Estructural del Instituto de Biología Molecular de Barcelona
(IBMB-CSIC)

Director de la tesis

Tutora de la tesis

Doctoranda

Joaquim Roca Bosch

Gemma Marfany Nadal

Joana Segura Martinez

Barcelona, Junio 2020

“Yo no estoy loco, mi realidad es diferente a la tuya”

El Gato de Cheshire

A mis padres

A mi hermana

Agradecimientos

Enfrentarse a esta página de agradecimientos como punto y final a una etapa tan compleja se hace difícil. Aparecen todas las emociones habidas y por haber, incertidumbres, miedos, esperanzas...

Primero de todo quiero agradecer a Joaquim por la confianza. Por permitirme formar parte de su laboratorio, hacerme descubrir el gusto por la ciencia bien hecha y enseñarme a prestar atención a todos los detalles, por muy pequeños que sean. Espero haber aprendido todo lo que me mostraba. *Moltes gràcies per guiar-me en aquesta etapa tan complexa i enriquidora.* Y a Gemma, como tutora, por acompañarme en todo momento y resolver todas mis dudas.

Dentro del laboratorio he tenido grandes compañeros que han pasado a ser indispensables de vida. Gracias a vosotros he aprendido la importancia de la colaboración y el trabajo en equipo, indispensable en ciencia. Gracias a Ofelia, por enseñarme todo dentro del laboratorio, por el apoyo, los consejos y hacerme sentir tan querida. A Belén, por resolver todas mis dudas experimentales (que no han sido pocas), por las conversaciones para solucionar el mundo y la vida, por ser mi “madre” dentro del laboratorio. A Toni, empezamos juntos en esto y te convertiste en mi gran apoyo, evolucionamos y aprendimos juntos, gracias por quedarte. A Sílvia, una gran compañera tanto dentro como fuera del laboratorio. Gracias por los cotilleos y las risas. A Xavi, por hacernos sentir siempre especiales. Y a Ricky, por su apoyo y disposición para ayudar siempre. Estas gracias las hago extensibles a todo el PCB e IBMB-CSIC, pasado y presente, siento que cada uno ocupa su lugar, como en una gran familia. Gracias a Antonia, Iracema, Luca, David, Raúl, Sara, Miren, Alexis, Marcos, Mònica, Pablo, María, Arkadiusz, Elena... Gracias por ayudarme a crecer tanto personal como científicamente. Me ha encantado conocerlos a todos.

Del PCB también me llevo un gran grupo de amigas y compañeras en *Noies al Lab*: Celia, Lada, Tere, Daniela, Zoë... Gracias por enseñarme a ser crítica y ver la ciencia y el mundo con las gafas de color violeta.

Que yo haya llegado hasta aquí no podía ser sin mi familia: Los Martínez, mis tíos, mis abuelos... Pero en especial, y esencialmente, por mis padres y mi hermana. Por el esfuerzo descomunal que mis padres hicieron y hacen por nosotras, para darnos todo lo que a ellos se les había negado. Por enseñarnos a ser fuertes y valientes, pero también a querernos si nos rompemos. Gracias, ojalá fuerais eternos.

Gracias a Laura, por ser mi compañera de vida, tú también deberías ser eterna. Junto a Erica y Núria, y a riesgo de sonar convencional, me habéis enseñado qué es la amistad.

A Hector, por ayudarme a buscar el equilibrio y retarme a nuevos desafíos. Gracias por querer quedarte y acompañarme en la vida. Llegar hasta aquí, y poder escribir estas páginas, también ha sido gracias a ti.

En estas últimas semanas pensaba sobre qué era lo que me había hecho querer ser científica. Gracias a mi tío Joaquín, *el tiet*, empecé a sentir curiosidad por las cosas, los animales y las plantas. Y ya, en el instituto, cuando descubrí la biología molecular, la genética y la química de la mano de profesoras como Guillermina o Mercedes, tuve clara mi carrera. En el laboratorio de Genómica Animal de la UAB con Aurora como mi tutora de máster, pude empezar mi camino por los laboratorios y la investigación. Muchas gracias por confiar en mí, por la paciencia y por ser una gran mentora. Te estaré siempre agradecida. Laia, Anna, Saray, Marina, Fran, Pablo... hicieron ese camino mucho más fácil, seguro que no habría sido lo mismo sin ellos.

Y aquí estoy, acabando el doctorado al "final" de una pandemia, esperando que sea un punto y seguido en mi aventura científica.

Gracias a todas por formar parte de ella, espero que estéis siempre.

Joana

sinopsis

SINOPSIS

El empaquetamiento del ADN en fibras de cromatina implica una gran cantidad de deformaciones geométricas de la doble hélice (enrollamiento y superenrollamiento del ADN) y la formación de una jerarquía de asas o dominios de distinto tamaño. Este empaquetamiento conlleva severas restricciones topológicas, las cuales juegan un papel determinante en los procesos metabólicos del ADN. Sin embargo, la arquitectura fina de la cromatina en 3D/4D, así como la trayectoria espacial y la topología de las cadenas de ADN a través de cada elemento individual de la cromatina, permanecen como aspectos enigmáticos y poco explorados. Ello es debido tanto a limitaciones metodológicas como a su dificultad conceptual, que hacen de la topología del ADN un área de estudio casi desconocida para gran parte de la comunidad científica.

En los años 80, diversos estudios experimentales establecieron que la unidad estructural de la cromatina, el nucleosoma, estabilizaba una diferencia del número de enlace (ΔLk) igual a -1.0/por nucleosoma. Este valor, sin embargo, suponía una incongruencia respecto al valor teórico estimado a partir de otros estudios estructurales de los nucleosomas, que concluían que el valor ΔLk debería ser > -1.0 y cercano a -2. Esta incongruencia generó la llamada “Paradoja del número de enlace del ADN nucleosomal”, que ha estado vigente hasta la actualidad.

En esta tesis se ha realizado un minucioso análisis de la topología del ADN nucleosomal para determinar con precisión el valor ΔLk . Los resultados presentados demuestran que los nucleosomas estabilizan un valor medio de $\Delta Lk = -1.26$. Este valor, junto a las correcciones de los valores teóricos derivados de los estudios estructurales previos, soluciona la paradoja del número de enlace del ADN nucleosomal. La aproximación metodológica, utilizando pequeños minicromosomas circulares para el análisis topológico de elementos de la cromatina *in vivo*, hace del presente trabajo una apuesta novedosa y original en el campo de la topología.

ABSTRACT

Packaging of DNA molecules into chromatin imposes severe geometrical deformations and topological constraints in the double helix (DNA wrapping, looping and supercoiling). These restrictions play an important role in the regulation of genome activities. Whereas the 3D/4D architecture of chromatin is being uncovered down to Kb resolution, the local path and the topology of the DNA strands constrained by individual chromatin elements remains elusive due to methodological limitations and conceptual complexity. In the '80s, it was established that the structural unit of chromatin, the nucleosome, constrains on average about one negative DNA supercoil (i.e. a DNA linking number difference, ΔLk , of $-1/$ per nucleosome). However, this ΔLk value was inconsistent with the theoretical value estimated from structural studies ($\Delta Lk \approx -2$). This incongruence between the theoretical and the experimental value produced the so-called "linking number paradox". In this thesis, we developed an approach to analyze the topology of DNA constrained by individual chromatin elements *in vivo*, so we could accurately determine the ΔLk value stabilized by individual nucleosomes. Our result indicates that nucleosomes stabilize an average value of $\Delta Lk = -1.26$. This value balances the twist ($\Delta Tw \approx +0.2$) and writhe ($\Delta Wr \approx -1.5$) deformations of nucleosomal DNA in terms of the equation $\Delta Lk = \Delta Tw + \Delta Wr$. Our finding could solve the long standing "linking number paradox" of nucleosomal DNA.

índice

SINOPSIS	1
ABSTRACT	5
INTRODUCCIÓN	11
1. ESTRUCTURA Y TOPOLOGÍA DEL ADN	13
1.1. Topología del ADN superenrollado	15
1.1.1. Número de enlace	15
1.1.2. <i>Lk</i> del ADN superenrollado	17
1.1.3. Número de torsión y número de enrollamiento	18
1.1.4. Conformación del ADN superenrollado	20
1.2. Topoisomerasas	23
1.2.1. Topoisomerasas Tipo IA	24
1.2.2. Topoisomerasas Tipo IB	26
1.2.3. Topoisomerasas Tipo II	28
1.3. Análisis del superenrollamiento del ADN mediante electroforesis	31
2. CROMATINA	34
2.1. Nucleosoma canónico	34
2.2. Variantes de las histonas	37
2.3. Modificaciones postraduccionales	38
2.4. Nucleosomas no-canónicos	39
2.5. Posición de los nucleosomas	41
2.6. Determinantes del posicionamiento nucleosomal	42
2.7. Dinámica conformacional de los nucleosomas	44
3. TOPOLOGÍA DEL ADN NUCLEOSOMAL	45
3.1. La paradoja del número de enlace	47

3.2.	Hipótesis para la resolución de la paradoja del número de enlace	47
3.2.1.	<i>Overtwisting</i> del ADN nucleosomal	47
3.2.2.	Contribución del ADN <i>linker</i>	49
3.3.	Topología de los nucleosomas atípicos: el centrómero puntual de <i>Saccharomyces cerevisiae</i>	50
4.	HIPÓTESIS DE PARTIDA	53
5.	ESTRATEGIA EXPERIMENTAL	55
	OBJETIVOS	61
	PUBLICACIONES	65
	Informe del director de tesis	67
	Resumen del primer artículo	73
	Artículo #1	75
	Resumen del segundo artículo	97
	Artículo #2	99
	DISCUSIÓN GENERAL	115
	CONCLUSIONES	127
	BIBLIOGRAFÍA	131

introducción

1. ESTRUCTURA Y TOPOLOGÍA DEL ADN

Los estudios de difracción de rayos X sobre fibras de ADN realizados por Rosalind Franklin permitieron dibujar su estructura tridimensional (Franklin & Gosling, 1953; Watson & Crick, 1953). Se observó una estructura consistente en una doble hélice de cadenas polinucleotídicas antiparalelas con las bases nitrogenadas de los nucleótidos emparejadas de forma específica mediante puentes de hidrógeno (A=T, C≡G). La doble hélice gira en sentido dextrógiro sobre un eje central, formando un interior hidrofóbico de bases nitrogenadas y un esqueleto hidrofílico exterior de azúcar-fosfato (Figura I1). Esta estructura helicoidal le confiere estabilidad pero produce una singularidad topológica: la separación de las dos hebras implica que éstas deben rotar y desenrollarse. Debido a la gran longitud de las moléculas del ADN y su elevado plegamiento en el interior del núcleo celular son necesarios mecanismos de desenrollamiento para llevar a cabo, de manera efectiva, los procesos metabólicos del ADN como su replicación y transcripción.

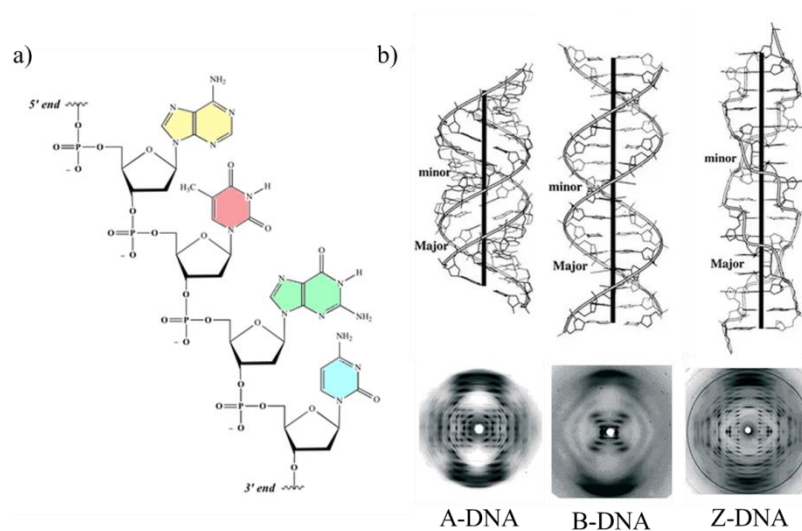


Figura I1. Estructura del ADN. (a) Estructura primaria del ADN. (b) Los estudios de difracción de rayos X sobre fibras y cristales de ADN revelaron variaciones en la estructura de la fibra en función de la composición de bases y las condiciones físicas en las que se encuentra: A-DNA, B-DNA y Z-DNA (Dickerson *et al.*, 1982).

A principios de los años 60, fueron descubiertas moléculas circulares de ADN en virus. En el trabajo publicado por Vinograd *et al.*, en 1965, se observó que el ADN vírico presentaba dos componentes con diferentes coeficientes de sedimentación aun siendo del mismo peso molecular. El componente I tenía un coeficiente de sedimentación mayor, es decir, era más compacto que el componente II. En el caso del componente I, las dos cadenas de ADN se mantenían unidas después de la desnaturalización, lo que sugería que se trataba de moléculas de ADN circular covalentemente cerradas. Se postuló entonces el componente II como la forma lineal de la misma molécula. Sin embargo, el componente II se generaba tras producir un solo corte (mediante ADNasa I) en una de las cadenas del componente I, poniendo en duda que el componente II fueran moléculas lineales. Posteriormente, mediante micrografía electrónica, se observó que tanto las moléculas del componente I como del II eran moléculas circulares pero con diferente grado de superenrollamiento. Las moléculas del componente I se encontraban superenrolladas, de la misma manera que se retuerce una goma elástica cuando la rotamos por un extremo. Y, las moléculas del componente II eran anillos sin tensión. Esto aclaraba la conversión del componente I en el componente II al producirse un corte (*nick*) en una de las cadenas. El corte permitía disipar la tensión helicoidal (superenrollamiento) de la doble hélice del ADN.

El superenrollamiento del ADN se ha observado en casi todos los sistemas biológicos. La tensión helicoidal del ADN facilita o dificulta la apertura de la doble hélice y desencadena conformaciones estructurales que alteran las interacciones y la mecánica de las moléculas de ADN.

1.1 Topología del ADN superenrollado

La topología del ADN superenrollado se define matemáticamente mediante tres parámetros principales: el número de enlace o *linking number* (Lk), el número de torsión o *twist number* (Tw) y el número de enrollamiento o *writhe number* (Wr).

1.1.1 Número de enlace

El valor de número de enlace o *Linking number* (Lk) es el número de veces que una curva (cadena o hebra) se enrolla o se enlaza alrededor de otra. En el caso de un ADN circular covalentemente cerrado, el valor de Lk es un número entero y, por convención, positivo al referirse al ADN dextrógiro. Así, el Lk es una propiedad estrictamente topológica e independiente de la geometría del ADN. Sólo puede variar cuando, tras la rotura de una o las dos hebras, se libera o añade una vuelta helicoidal entre ellas, antes de volver éstas a unirse (Figura 12). El Lk aumenta proporcionalmente con la longitud del ADN. Su valor corresponde al cociente entre la longitud (N) dividido por el número de pares de bases (pb) por vuelta de la doble hélice (h , *helical repeat*):

$$Lk = \frac{N}{h}$$

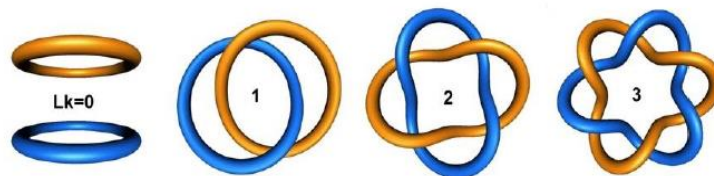


Figura 12. Número de enlace (Lk). Representación de los valores del número de enlace (Lk) entre dos curvas cerradas.

Box1: Valor de h o *helical repeat*

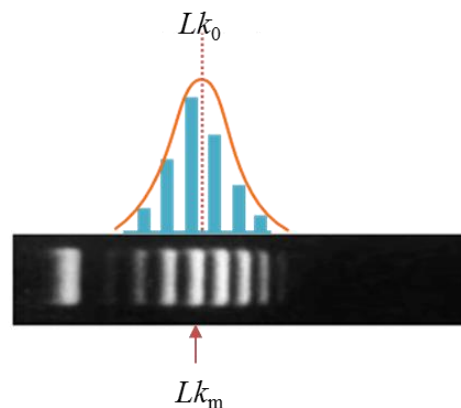
El valor de h es el número de pares de bases necesarias para que la hebra de ADN complete una vuelta helicoidal alrededor de la otra hebra. Este valor depende de las condiciones del medio en las que se encuentra el ADN (concentración de sales, pH y temperatura), además de las posibles

interacciones con moléculas presentes en el medio que actúan como ligandos. En una disolución con condiciones fijadas como fisiológicas (NaCl al 0.2M, pH 7 y 37°C), un ADN desnudo y sin tensión presenta un valor de h denominado $h^0 = 10.5$ pb/vuelta.

Debido a las variabilidades del medio, el cálculo de h se hace complejo. Una aproximación experimental se realiza mediante el análisis de la periodicidad en los cortes realizados por la ADNasa I cuando el ADN se encuentra unido a una superficie. Pero esta unión también afecta al valor de h . El cálculo más ajustado se realizó en disolución, añadiendo bases de una en una a un anillo de ADN. Tras cada adición se analizaba el Lk del anillo mediante electroforesis. Tras añadir 21 bases, se observó que el valor Lk había aumentado exactamente en 2 unidades (Wang, 1979).

Debido a las oscilaciones termodinámicas a las que se encuentra sometido el ADN en disolución, el valor de h de las moléculas fluctúa con centro en h^0 (10.5 pb). Por tanto, al circularizar moléculas de ADN lineal, se genera una distribución Gaussiana de valores de Lk (distribución de Boltzman) con centro en Lk_0 (Wang, 1979; Peck & Wang, 1981). La distribución de valores Lk se resuelve mediante electroforesis en gel de agarosa, donde se observa una escalera de topoisómeros correspondiente a las moléculas con diferente Lk . Al tratarse de una distribución Gaussiana de valores, su centro o Lk_0 , puede no ser un valor entero y no coincidir con un topoisómero concreto. Esto ocurre ya que, al circularizar un ADN lineal, los extremos no siempre quedan exactamente alineados para el enlace 5'-3', por lo que se requerirá un pequeño giro (torsión) para ligarlos. Cuando Lk_0 no coincide con un Lk entero concreto, se denomina Lk_m al valor íntegro de la distribución más cercano a Lk_0 (Figura 13).

Figura 13. Distribución de topoisómeros en gel de electroforesis. La circularización de moléculas de ADN genera una distribución Gaussiana (distribución de Boltzman) con centro en Lk_0 . El valor íntegro más cercano a Lk_0 se denomina Lk_m y corresponde al topoisómero con la mínima tensión requerida para su circularización.



1.1.2 Lk del ADN superenrollado

El incremento o reducción del número de vueltas helicoidales antes de circularizar una molécula de ADN da lugar a moléculas con un Lk mayor o menor que Lk_0 , lo que se traduce en la introducción de tensión helicoidal. Esta diferencia de Lk respecto al estado relajado Lk_0 , $Lk-Lk_0$, es conocida como diferencia de número de enlace, o ΔLk , la cual puede ser positiva o negativa y es indicativo del grado de superenrollamiento del ADN (Figura I4).

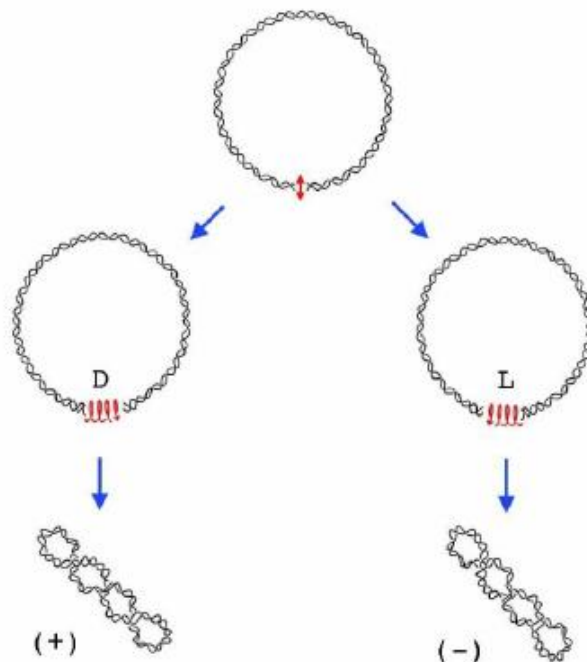


Figura I4. Superenrollamiento por variación en el valor de Lk . Las variaciones del número de enlace antes de la circularización de una molécula de ADN se traducen en tensión helicoidal que genera el superenrollamiento del ADN. Los giros dextrógiros (D) incrementan el valor de Lk dando lugar a superenrollamiento positivo (+), mientras que los giros levógiros (L) reducen el Lk dando lugar a superenrollamiento negativo (-).

La tensión asociada a ΔLk varía según la longitud de la molécula. Así, ΔLk puede relativizarse al tamaño del ADN circular al dividir por Lk_0 (el cual es proporcional al

número de bases). Este valor se conoce como diferencia específica de enlace o densidad superhelicoidal (σ):

$$\sigma = \frac{\Delta Lk}{Lk_0}$$

Box2: Densidad de superenrollamiento in vivo:

Tanto en organismos procariotas como eucariotas, la densidad de superenrollamiento del ADN intracelular tiene un valor cercano a -0.06 (Bauer, 1978). Es decir, el Lk de la doble hélice presenta un déficit del 6% respecto al valor de Lk_0 correspondiente al ADN desnudo y relajado. De forma excepcional, en eubacterias y arqueas termófilas extremas, el ADN se estabiliza in vivo adoptando un mayor valor de Lk debido a que se encuentra a altas temperaturas (>80°C). Al extraer los plásmidos de estas bacterias y analizarlos a temperatura ambiente, se observa que el ADN presenta un superenrollamiento positivo (Lopez-Garcia & Forterre, 1997).

La energía asociada al superenrollamiento produce cambios geométricos y conformacionales de la molécula facilitando o impidiendo transiciones estructurales. Estos cambios pueden ser condicionantes para la interacción del ADN con proteínas y otros ligandos.

1.1.3 Número de torsión y número de enrollamiento

Los cambios producidos en las moléculas de ADN debido a variaciones en ΔLk generan deformaciones geométricas como respuesta. Éstas se pueden describir mediante dos propiedades: el número de torsión y el número de enrollamiento.

El número de torsión (Tw) describe el número de giros de las hebras del ADN alrededor del eje central de la doble hélice. Por convención, de la misma manera que el valor de Lk , el Tw tiene un valor positivo en ADN dextrógiro. Por otro lado, el número de enrollamiento (Wr) describe la trayectoria espacial del eje de la doble hélice y varía en sentido positivo o negativo cada vez que el eje central del ADN se cruza consigo mismo al proyectarse sobre un plano. Los valores de Tw y Wr son difíciles de calcular con exactitud cuando hablamos de un dominio de ADN. Sin embargo, estos dos valores geométricos están íntimamente ligados al ser su sumatorio equivalente al valor de Lk (White, 1969, Fuller, 1978):

$$Lk = Tw + Wr$$

Esta simple ecuación implica que, mientras el valor de Lk sea constante (en dominios de ADN con extremos anclados o en moléculas circulares covalentemente cerradas), cualquier cambio en Tw vendrá acompañado por un cambio de la misma magnitud, pero de signo contrario, en forma de Wr y viceversa. En ausencia de tensión helicoidal ($Lk \approx Lk_0$), la doble hélice de ADN no se retuerce ($Wr \approx 0$) y su valor de Tw coincide, por tanto, con el valor de Lk ($Lk = Tw$). En presencia de tensión helicoidal ($Lk \neq Lk_0$), la doble hélice de ADN cambia su geometría alterando sus valores de Tw y Wr con una repartición aproximada del 30 y 70% respectivamente. De la misma forma, con la ecuación queda implícito que, un cambio de Lk (mediante corte y posterior ligación de la molécula) se manifestará de forma geométrica como un cambio en forma de Tw y/o de Wr (Vologodskii & Cozzarelli, 1994) (Figura 15):

$$\Delta Lk = \Delta Tw + \Delta Wr$$

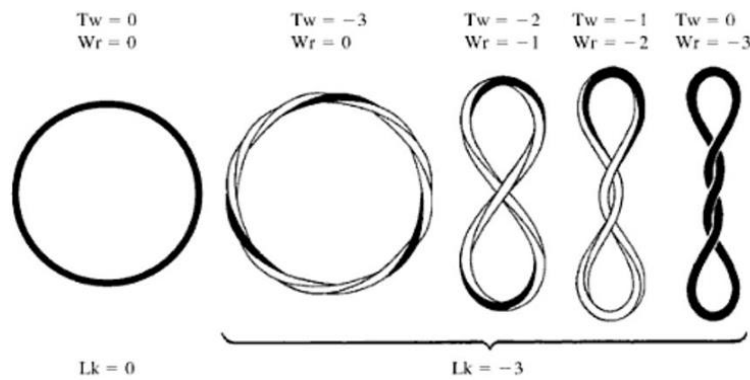


Figura 15. Relaciones topológicas. Interconversión entre los valores de Tw y Wr sin variación del valor de Lk en un anillo cerrado.

Los cambios en forma de Tw implican una apertura o cierre de la doble hélice, es decir, un aumento o disminución del número de pares de bases por vuelta. Así, un ΔLk negativo puede generar la apertura de la doble hélice y por tanto, la disminución del valor de Tw . En cambio, con un ΔLk positivo, un aumento en el valor de Tw se puede producir mediante la extrusión de algunas bases hacia el exterior de la doble hélice. En cuanto a las

variaciones geométricas en forma de Wr , la molécula de ADN adopta un superenrollamiento espacial -propriadamente dicho- (*supercoiling*), que puede ser positivo o negativo y dar lugar a conformaciones del superenrollamiento de tipo solenoide o plectonómico.

1.1.4 Conformaciones del ADN superenrollado

La deformación en solenoide, o toroidal, se caracteriza por el giro del dúplex alrededor de un cilindro (o toro) imaginario formando una superhélice. Normalmente, es necesaria la interacción del ADN con proteínas u otros ligandos para estabilizar dicha deformación, de la misma manera que el ADN se envuelve sobre los octámeros de histonas para formar los nucleosomas. La superhélice formada será dextrógira o levógira en función de si la densidad superhelicoidal (σ) es positiva o negativa, respectivamente (Figura I6a). Por el contrario, si no hay proteínas u otros ligandos para estabilizar la tensión helicoidal (T_h) del ADN, la energía libre asociada a las deformaciones en forma de ΔTw o de ΔWr provoca un plegamiento espontáneo de la molécula de ADN formando un plectonema. Esta deformación se caracteriza por el mutuo entrelazado de dos segmentos antiparalelos del dúplex formando una súper doble hélice (tal y como se retuerce una goma elástica con tensión helicoidal) (Adrian *et al.*, 1990; Bates & Maxwell, 2005). En función de si T_h es positiva ($\sigma > 0$) o negativa ($\sigma < 0$), la súper doble hélice del plectonema es levógira o dextrógira, respectivamente, es decir, lo contrario a lo que ocurre en el solenoide (Figura I6b).

Además de las conformaciones superenrolladas, el ADN puede presentarse *in vivo* formando nudos y encadenamientos donde una o más moléculas se encuentran entrelazadas debido, en parte, a la alta densidad de material genético dentro del núcleo (Figura I6c y I6d).

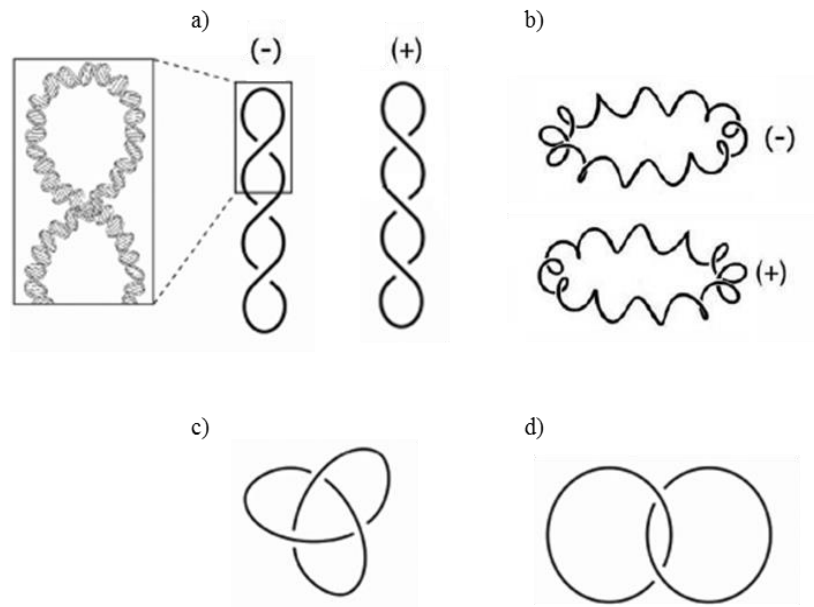


Figura 16. Conformaciones del ADN. El ADN superenrollado puede adoptar conformaciones plectonémicas (a) o solenoidales (b) si se trata de ADN desnudo o puede interactuar con proteínas, respectivamente. Los nudos (c) y encadenamientos (d) son invariantes topológicas las cuales requieren de rotura de la doble hélice para su resolución.

Entendemos como nudo topológico cuando un dominio cerrado (ej. anillo de ADN) entrelaza consigo mismo de forma que sus intersecciones sólo pueden eliminarse al cortar segmentos de la curva. La complejidad del nudo vendrá dada por el número de estas intersecciones. Así, el nudo más simple es el formado por 3 intersecciones, denominado 3_1 o trébol, seguido del nudo de 4 cruces (4_1). A medida que aumenta el número de cruces, aumenta la complejidad, apareciendo diferentes formas en las que se pueden realizar dichos cruces. En el caso del nudo con 5 cruces, encontramos 2 formas (5_1 y 5_2), el nudo de 6 cruces presenta 3 formas posibles, etc. Así, la complejidad aumenta exponencialmente llegando a miles de formas de nudos distintos en el caso de anillos con más de 10 intersecciones (Figura 17).

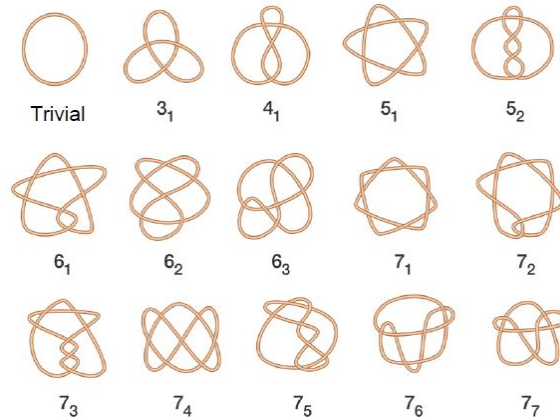


Figura 17. Nudos. Representación de nudos de complejidad creciente en un dominio circular cerrado.

Las primeras observaciones experimentales de nudos de ADN se realizaron en bacteriófagos (Liu *et al.*, 1981). En estos casos, los extremos cohesivos del ADN lineal vírico propician la circularización accidental del ADN en el interior de las pequeñas cápsides de los fagos y, por tanto, la formación de nudos (Arsuaga *et al.*, 2002). Posteriormente, se encontró ADN anudado en bacterias (Shishido *et al.*, 1987) y, más recientemente, se ha observado por primera vez en células eucariotas. En este trabajo, realizado en nuestro laboratorio, los nudos de ADN aparecen con una probabilidad del 2-3% en regiones de 5 kilobases (kb) (Valdés *et al.*, 2018).

Por otro lado, los encadenamientos implican el enlace topológico entre diferentes moléculas de ADN. A diferencia de los nudos, los encadenamientos se han observado en una gran variedad de sistemas biológicos (Bates & Maxwell, 2005). Ello es debido a que, durante la replicación del ADN, las moléculas nacientes de ADN quedan siempre encadenadas. Ya que estos encadenamientos impedirían la segregación de los cromosomas en el momento de la división celular, son necesarios cortes transitorios del ADN, realizados por las enzimas llamadas topoisomerasas (Wang, 1996).

1.2 Topoisomerasas

Las topoisomerasas son una familia de enzimas que modifican la topología del ADN. Todas las topoisomerasas comparten el mismo principio de acción que implica el corte transitorio de una o las dos hebras de ADN formando un intermediario covalente entre el grupo hidroxilo de una tirosina y un fosfato del ADN. La topoisomerasa permite entonces el paso de otro segmento de ADN entre los extremos escindidos y, finalmente, repara la escisión realizada (Figura I8).

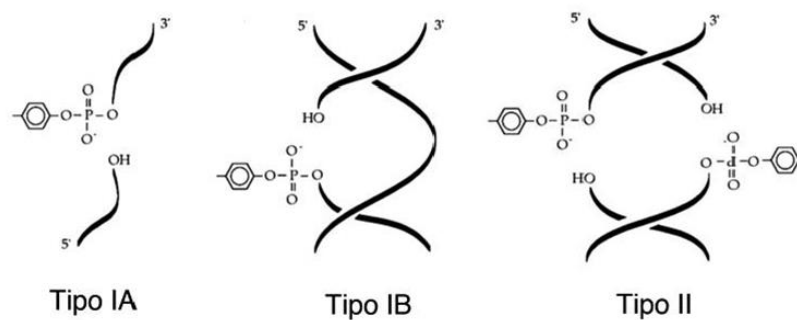


Figura I8. Clasificación de las topoisomerasas según su principio de acción. Las topoisomerasas generan cortes transitorios en una hebra (Tipo IA y Tipo IB) o las dos hebras (Tipo II) del ADN. A través del corte generado permiten el paso de otro segmento del ADN, alterando la topología.

Desde su descubrimiento en 1971 (Wang, 1971), las topoisomerasas se han encontrado en todo tipo de células tanto procariontas como eucariotas, además de diversos virus bacterianos que codifican sus propias topoisomerasas. Son enzimas muy conservadas evolutivamente y se clasifican en tres familias (Figura I9). Las topoisomerasas de tipo I, que producen cortes en una hebra, y las topoisomerasas de tipo II, que actúan escindiendo las dos hebras de la doble hélice. A su vez, las de tipo I se subdividen en topoisomerasas tipo IA y tipo IB, en función de su mecanismo de acción y la polaridad (3' o 5') del enlace covalente entre la enzima y el ADN (Figura I8 y Tabla I1).

Familia	Enzima	Procedencia
Tipo IA	Topoisomerasa I y III	<i>Escherichia coli</i>
	Topoisomerasa III	<i>Saccharomyces cerevisiae</i>
	Topoisomerasa III α y III β	<i>Drosophilla</i>
	Topoisomerasa III α y III β	Mamíferos
Tipo IB	Topoisomerasa I	Eucariotas
	Topoisomerasa I	Mitocondria
	Topoisomerasa I	<i>Pox virus</i>
Tipo II	Girasa	<i>Escherichia coli</i>
	Topoisomerasa IV	<i>Escherichia coli</i>
	Topoisomerasa	Fago T4
	Topoisomerasa II	<i>Saccharomyces cerevisiae</i>
	Topoisomerasa II	<i>Drosophilla</i>
Topoisomerasa II α y II β	Mamíferos	

Tabla I1. Enzimas representativos de cada familia de topoisomerasas.

1.2.1 Topoisomerasas Tipo IA

Las topoisomerasas de tipo IA cortan hebras simples de ADN formando el intermediario covalente entre la tirosina del centro activo con el extremo 5'-fosfato del ADN (Wang, 1996). A través del corte, permiten el paso de otra cadena simple de ADN (Figura I9). No requieren cofactores energéticos, por lo que producen cambios topológicos mediante la energía libre de los segmentos de ADN involucrados.

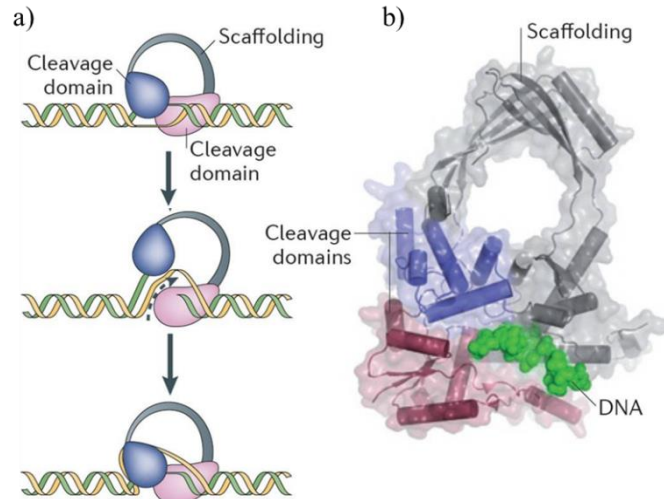


Figura I9. Topoisomerasas Tipo IA. (a) Esquema del mecanismo de acción de las topoisomerasas Tipo IA. La enzima se une a una hebra de ADN en la que genera un corte a través del cual pasa otra hebra hasta la cavidad central de la enzima. Mediante este proceso permite reducir la tensión helicoidal negativa del ADN. (b) Estructura de la topoisomerasa III de *E. coli* unida a un segmento de ssDNA (Vos *et al.*, 2011).

Las topoisomerasas tipo IA relajan la tensión helicoidal negativa ($Lk < Lk_0$), la cual permite la separación parcial de las hebras del ADN favoreciendo así la acción de la enzima. Estas topoisomerasas también pueden encadenar/desencadenar y anudar/desanudar anillos o dominios de cadena simple (ssDNA) o bien de cadena doble (dsDNA) si contienen segmentos de cadena simple (*nicks o gaps*) (Figura I10). La enzima representativa de este grupo es la topoisomerasa I de *E.coli*, así como las topoisomerasas III presentes en bacterias y eucariotas (Wang, 1971; Wallis *et al.*, 1989).

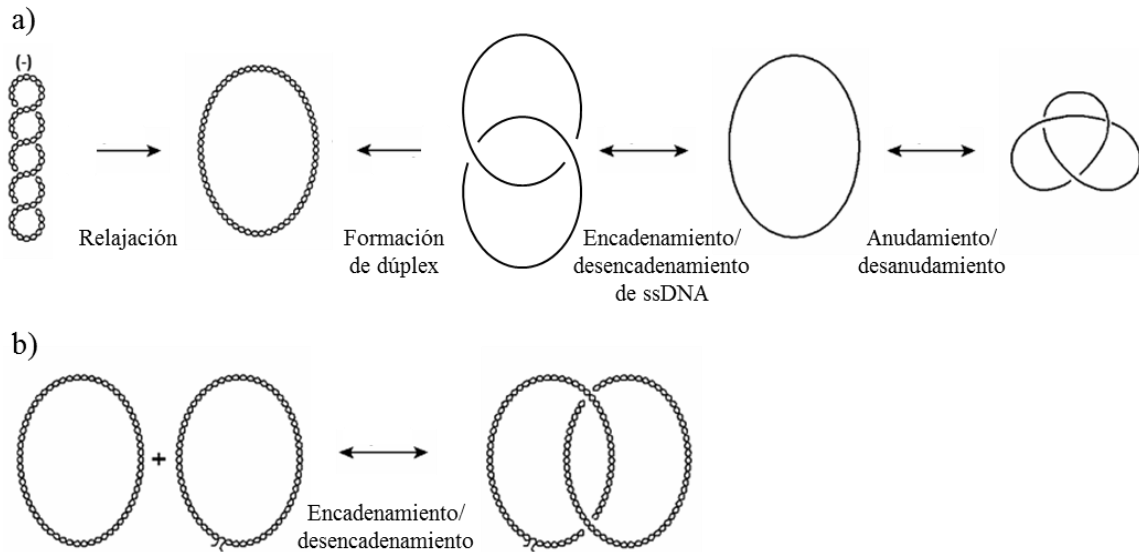


Figura I10. Conversiones topológicas catalizadas por las Topoisomerasas Tipo IA. Las topoisomerasas IA permiten reducir la tensión helicoidal negativa del ADN además de formar nudos y encadenamientos a partir de ssDNA (a) o bien de dsDNA si contiene *nicks* o *gaps* (b).

1.2.2 Topoisomerasas Tipo IB

Las topoisomerasas tipo IB, a diferencia de las tipo IA, actúan sobre moléculas de dsDNA y generan un *nick* transitorio sólo en una de las hebras. En este caso, el intermediario covalente tiene lugar entre la tirosina y el extremo 3'-fosfato del ADN. El extremo 5'-OH sólo interacciona de forma iónica por lo que puede rotar alrededor de la hebra no escindida. Tras una o varias rotaciones, la enzima vuelve a sellar la cadena cortada (Figura I11).

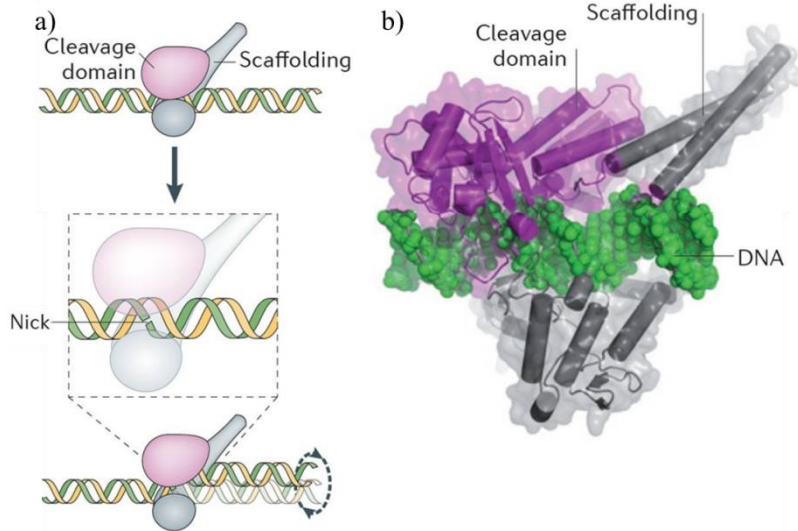


Figura I11. Topoisomerasas tipo IB. (a) Esquema del mecanismo de acción de las topoisomerasas tipo IB. La topoisomerasa genera un *nick* transitorio en una de las cadenas del dúplex de ADN formando una unión covalente con el extremo 3'-fosfato. El extremo 5'-OH libre puede rotar liberando la tensión antes de la resolución del *nick*. (b) Estructura de la topoisomerasa IB humana (Vos *et al.*, 2011).

Este mecanismo permite relajar de forma muy eficiente las moléculas de ADN con tensión helicoidal tanto positiva como negativa dando lugar a distribuciones de Lk centradas en Lk_0 (Champoux *et al.*, 1990) (Figura I12).

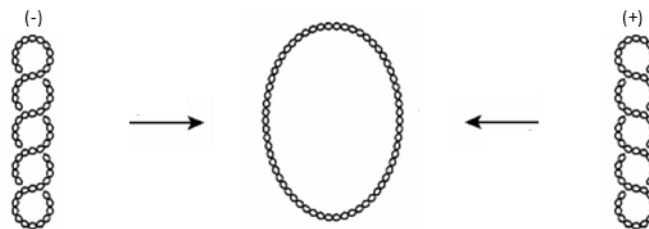


Figura I12. Conversiones topológicas catalizadas por las Topoisomerasas Tipo IB. Las topoisomerasas tipo IB permiten relajar completamente tanto la tensión helicoidal positiva como negativa del ADN, generando distribuciones térmicas de Lk centradas en Lk_0 .

La enzima representativa de este grupo es la topoisomerasa I presente en todos los eucariotas. Siendo su extremo C-terminal, donde se encuentra la tirosina activa, un

dominio muy conservado evolutivamente (Eng *et al.*, 1989; Lynn *et al.*, 1989). En esta familia también se encuentra la topoisomerasa I de *Pox virus* (Vaccinia virus).

1.2.3 Topoisomerasas Tipo II

Las topoisomerasas de tipo II son homodímeros funcionales que permiten el transporte de un segmento de ADN de doble cadena a través de una escisión realizada en otro segmento de doble cadena (Hsieh, 1990). Además, a diferencia de las de tipo I, utilizan ATP como cofactor energético (Wang, 1996).

Estas enzimas están formadas por 3 dominios muy conservados evolutivamente: dominio ATPasa o *N-gate*, dominio de unión-escisión del ADN o *DNA-gate* y dominio pivote o *C-gate* junto con un dominio C-terminal menos conservado (Corbet & Berger, 2004). La actividad enzimática se inicia mediante la unión del segmento de ADN que va a ser escindido (segmento G) con el dominio *DNA-gate*. Durante la escisión, los extremos 5'-fosfato quedan unidos covalentemente con los residuos tirosina activos del *DNA-gate*, mientras que los extremos 3'-OH quedan sujetos mediante uniones no covalentes que evitan su libre rotación. La unión de ATP a la enzima provoca un cambio conformacional con el cierre del dominio *N-gate*, lo que permite que otro segmento de ADN (segmento T) pueda ser capturado hacia el interior de la enzima y empujado a través de la escisión del segmento G (Roca & Wang, 1992). Tras cruzar por el segmento G, el segmento T sale por el lado opuesto a su entrada a través del dominio *C-gate* (Roca & Wang, 1994; Roca *et al.*, 1996). En ocasiones, y de forma alternativa, el segmento T puede retroceder y salir de nuevo por el dominio *N-gate*. En este caso, no se produciría ningún cambio topológico (Martínez- García *et al.*, 2014) (Figura I13).

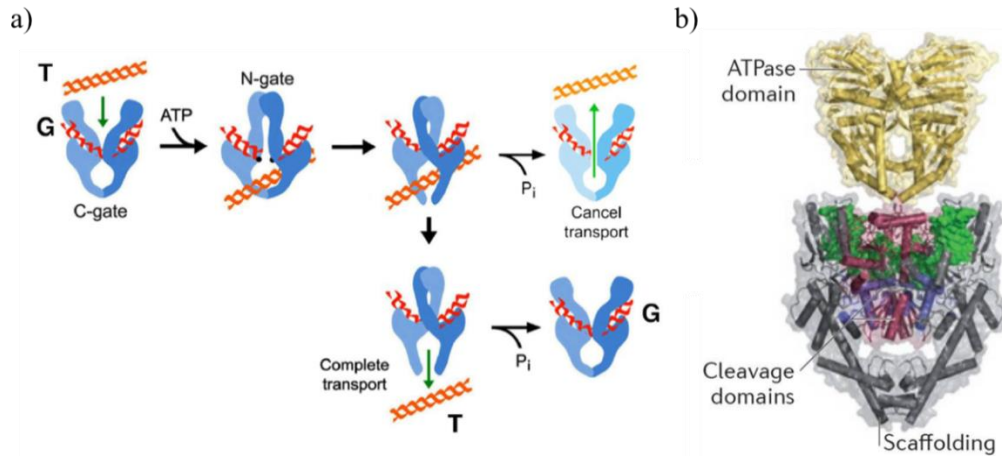


Figura I13. Topoisomerasas II. (a) Esquema del mecanismo de acción de las topoisomerasas Tipo IIA. Una vez unida la enzima al segmento G (rojo), el cambio de conformación por la unión de ATP permite capturar el segmento T (naranja). El segmento T puede salir por el lado opuesto de la enzima y generar un cambio topológico o retroceder y deshacer su recorrido, cancelándose el transporte (Martínez-García *et al.*, 2014). (b) Estructura de la topoisomerasa IIA de levadura unida a ATP y al segmento G (verde) (Vos *et al.*, 2011).

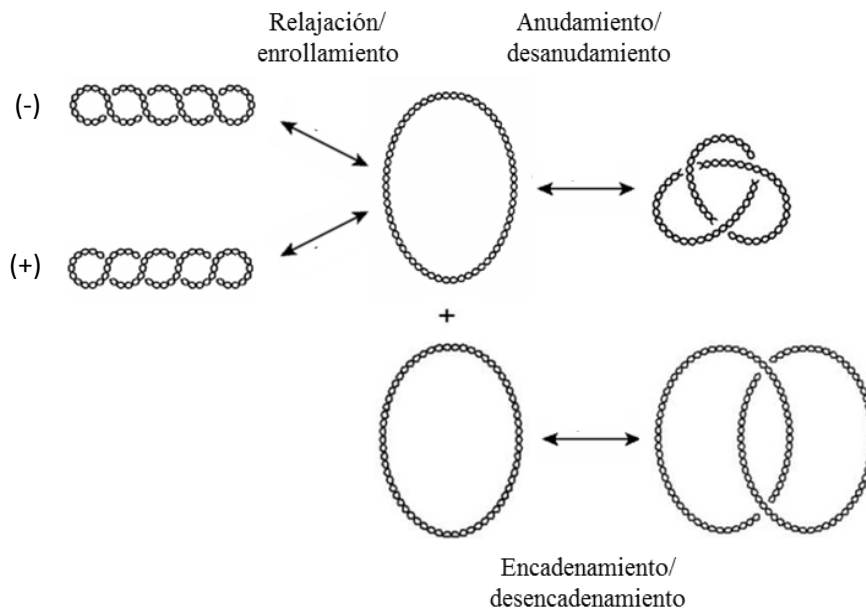


Figura I14. Conversiones topológicas catalizadas por las topoisomerasas tipo II. Según la localización de los segmentos G y T se pueden dar conversiones intramoleculares (relajación/enrollamiento y anudamiento/desanudamiento) o conversiones intermoleculares (encadenamiento/ desencadenamiento).

Si los segmentos implicados en esta acción se encuentran en la misma molécula de ADN, las topoisomerasas tipo II producen un cambio en el valor de Lk de 2 unidades, positivas o negativas según se invierta un cruce G/T negativo o positivo, respectivamente. Este cambio permite relajar tanto el superenrollamiento negativo como positivo del ADN. Estando los segmentos G y T en la misma molécula, las topoisomerasas tipo II también pueden formar o eliminar nudos en el ADN. En cambio, si los segmentos se encuentran en distintas moléculas, el resultado es la formación o eliminación de encadenamientos (Figura I14).

En esta familia está presente la topoisomerasa II, esencial en células eucariotas, la topoisomerasa IV y la girasa de ADN, las dos de eubacterias. La girasa de ADN es una enzima especial de tipo II (Gellert *et al.*, 1976). A diferencia del resto de topoisomerasas tipo II, presenta un dominio C-terminal que permite la interacción con segmentos G y T contiguos, de manera que el T forma un cruce (supercoil) positivo con el G (Figura I15). La inversión de este cruce de positivo a negativo, permite a la girasa reducir el valor de Lk y generar tensión helicoidal negativa (Figura I15).

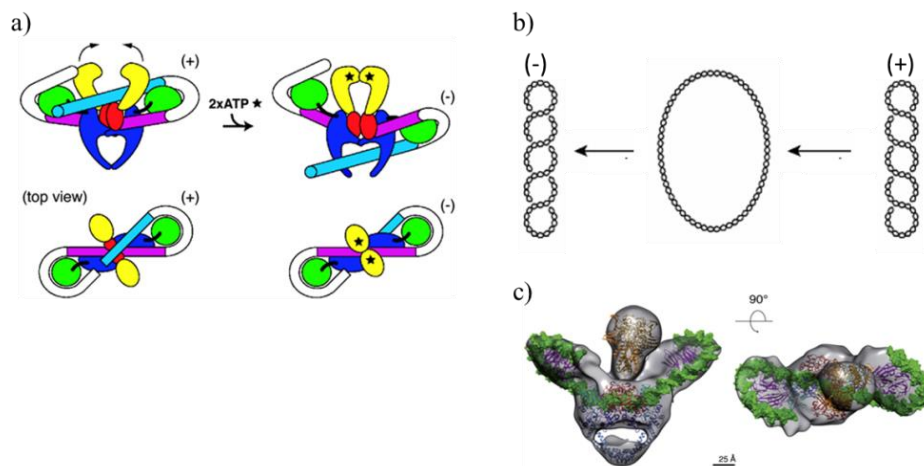


Figura I15. Girasa de ADN. La girasa de ADN interacciona de forma exclusiva con los cruces positivos y los invierte a negativos. (a) Mecanismo de acción de la girasa de ADN (Corbett *et al.*, 2004). (b) Conversiones topológicas catalizadas por la girasa. (c) Estructura de la girasa de ADN acomodando un fragmento de ADN (verde) (Papillon *et al.*, 2013).

1.3 Análisis del superenrollamiento del ADN mediante electroforesis

Los genomas consisten en largas moléculas de ADN plegadas formando bucles o *loops*, de decenas a cientos de kilobases (kb), que generan dominios topológicos. Cada dominio se comporta como un anillo circular de ADN donde no se permite la libre rotación de los extremos. Por este motivo, la topología del ADN se estudia en el laboratorio, generalmente, mediante la electroforesis de plásmidos o minicromosomas circulares como modelo.

La electroforesis en geles de agarosa permite separar moléculas de ADN según su longitud y/o grado de compactación, siendo las moléculas más cortas y/o más compactadas las que migran más rápido a través del gel. Al examinar un plásmido extraído de bacterias, se observan, normalmente, dos bandas. Una banda con mayor movilidad, correspondiente al anillo de ADN covalentemente cerrado con superenrollamiento negativo (muy compacto), y una banda con menor movilidad, correspondiente al anillo sin superenrollamiento (menos compacto) debido a la presencia de uno o varios *nicks*. La incubación del plásmido con topoisomerasa I produce su relajación (se aumenta su *Lk*), dando lugar a una distribución de topoisómeros observables como una escalera de bandas, donde cada una corresponde a un valor de *Lk* determinado (Figura I16).

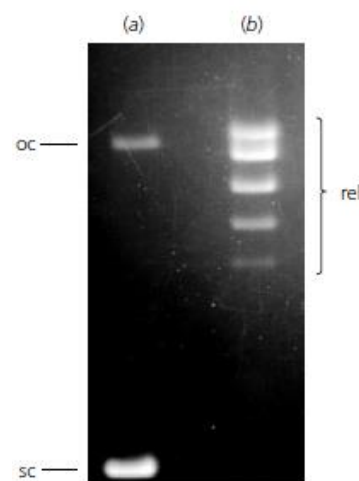


Figura I16. Electroforesis en gel de agarosa de un anillo de ADN. (a) El ADN plasmídico extraído de bacterias muestra dos poblaciones: anillo con superenrollamiento negativo (*sc*) y anillo con discontinuidad (*oc*). (b) Después de la incubación con topoisomerasas, el ADN plasmídico forma una distribución de topoisómeros (*rel*) (Bates & Maxwell, 2005).

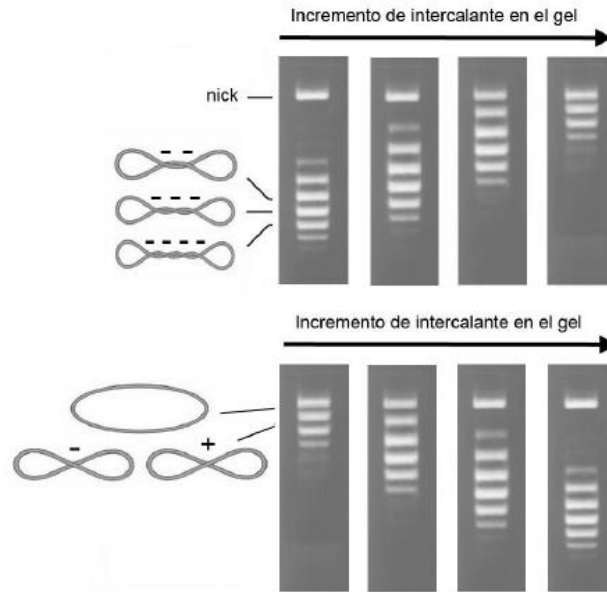


Figura I17. Efecto del intercalante en la movilidad electroforética de las distribuciones de Lk . En el panel superior se muestra cómo los topoisómeros con superenrollamiento negativo ($\Delta W_r < 0$) presentan una mayor velocidad de migración respecto al ADN relajado ($\Delta W_r = 0$) y el anillo con *nick*, debido a una mayor compactación. Al añadir el intercalante durante la electroforesis, se reduce el valor de T_w que es compensado mediante un incremento de W_r , disminuyendo la velocidad electroforética. En el panel inferior, los topoisómeros cercanos a Lk_0 , migran próximos al *nick*. Al añadir intercalante, se reduce el valor de T_w , por lo que se incrementa el valor de W_r . En consecuencia, adquieren una mayor velocidad electroforética quedando separados del *nick*.

Cada valor de Lk implica un valor de W_r y, por tanto, un grado de compactación distinto que explica las diferencias observadas en la movilidad electroforética. Sin embargo, la mera velocidad de migración no permite diferenciar entre un valor positivo o negativo de W_r , ya que el grado de compactación es el mismo en los dos casos. Para distinguir entre un valor positivo y negativo de W_r , es decir, entre superenrollamiento positivo y negativo, se añaden intercalantes de ADN durante la electroforesis. Estos intercalantes, como por ejemplo la cloroquina, alteran el valor de h al unirse al ADN y reducen el valor de T_w . Como se trata de moléculas de ADN circular covalentemente cerradas que mantienen su valor de Lk , esta reducción del valor de T_w se compensa con un incremento del valor de W_r . Un incremento de W_r descompacta las moléculas con superenrollamiento negativo, lo que disminuye su velocidad de migración en la electroforesis. En las moléculas relajadas y

las que ya tienen superenrollamiento positivo, un incremento de Wr las compacta, aumentando su velocidad de migración en la electroforesis. Este principio permite jugar experimentalmente con las concentraciones de intercalante para resolver las distribuciones de topoisómeros superenrollados positiva- o negativamente, según convenga (Figura I17).

En ocasiones, es difícil resolver las distribuciones de Lk mediante geles de una sola dimensión. En estos casos se pueden realizar electroforesis bidimensionales (2D). Estas electroforesis consisten en dos electroforesis sucesivas en direcciones ortogonales sobre un mismo gel (Figura I18). Las dos electroforesis se efectúan con concentraciones de intercalantes distintas, por lo que el valor de Wr (compactación) que adquiere cada topoisómero es diferente en cada dimensión (Hanai & Roca, 1999). Como resultado, los topoisómeros con distinto Lk se resuelven formando un arco donde se diferencian las moléculas con superenrollamiento positivo y superenrollamiento negativo, quedando éstas a su vez separadas de las moléculas lineales y moléculas con *nicks*.

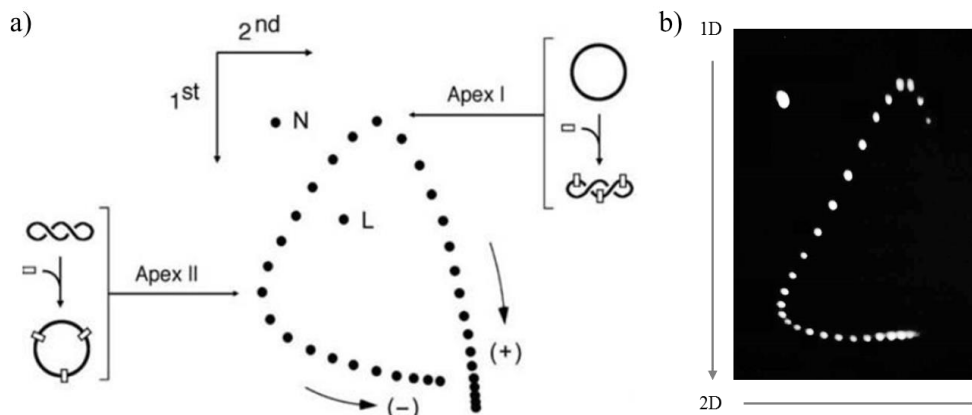


Figura I18. Separación de topoisómeros mediante electroforesis 2D. (a) Esquema de topoisómeros (puntos negros) en una electroforesis 2D. El Ápex I indica el topoisómero con menos Wr (menor compactación y velocidad) durante la primera dimensión. La unión de intercalante (rectángulos blancos), durante la segunda dimensión, aumenta el valor de Wr de este topoisómero de manera que presenta una velocidad mayor en la segunda dimensión, permitiendo su separación del *nick* (N). El Ápex II indica un topoisómero superenrollado negativamente que, al unirse con el intercalante, cambia a un valor de $Wr \approx 0$, disminuyendo su velocidad en la segunda dimensión. L indica los anillos lineales. (b) Electroforesis bidimensional de un plásmido de ADN.

2. CROMATINA

In vivo, las moléculas de ADN interaccionan con proteínas dando lugar a fibras de cromatina. Aunque la cromatina fue descrita en el siglo XIX por W. Flemming al observar una sustancia nuclear visible tras la tinción con colorantes básicos, no fue hasta mediados del siglo XX, con técnicas de obtención de cromatina soluble (Zubay & Doty, 1959), que se empezaron a dilucidar sus características biofísicas.

En células eucariotas, la cromatina es mayoritariamente un complejo de ADN con proteínas histonas que, junto a proteínas no-histonas y RNA, reduce en varios órdenes de magnitud el volumen que ocuparía el ADN celular desnudo. La cromatina es fundamental para regular la actividad génica ya que organiza el ADN nuclear en distintos territorios y dominios, establece regiones de mayor y menor accesibilidad y promueve contactos entre sitios distantes (Jiang & Pugh, 2009; Radman-Livaja & Rando, 2010). La unidad estructural básica de la cromatina eucariótica es el nucleosoma formado por un octámero de histonas, sobre el cual se enrolla el ADN (Figura I19).

En organismos procariotas, a diferencia de las células eucariotas, no hay un compartimento nuclear que confine el ADN. Aun así, el genoma bacteriano también se encuentra cromatinizado. Se han descrito proteínas análogas a las histonas asociadas al ADN bacteriano. Entre estas proteínas se encuentran las proteínas HU y H-NS, que empaquetan el ADN y participan en la regulación de la actividad génica bacteriana (McLeod & Johnson, 2001).

2.1 Nucleosoma canónico

El nucleosoma está extraordinariamente conservado a lo largo de la evolución. Su estructura presenta una simetría binaria y está formado por un núcleo o *core* proteico con ocho histonas sobre el que 147 pb de ADN completan cerca de 1.7 vueltas levóginas (Richmond *et al.*, 1984; Luger *et al.*, 1997). El octámero de histonas está compuesto por un tetrámero central con dos copias de las histonas H3 y H4 flanqueadas por dos dímeros

H2A y H2B [(H3+H4)₂ + (H2A+H2B)₁₊₁] (Figura I19). Las histonas presentan una región globular de aproximadamente 65 aminoácidos que forman tres α -hélices donde tienen lugar las interacciones de dimerización (Mariño-Ramírez *et al.*, 2006). Los extremos C- y N-terminal están menos estructurados y son más accesibles al proyectarse fuera del *core* nucleosomal (Davey *et al.*, 2002). Los nucleosomas se forman a lo largo de las moléculas de ADN de forma periódica cada 200 ± 40 pb, quedando separados entre ellos por segmentos de ADN desnudo (*ADN linker*) de 30 a 90 pb, según el tipo celular. De esta manera, bajo el microscopio electrónico, los nucleosomas se observan formando una fibra similar a un collar de cuentas de 10 nm de diámetro (*beads-on-a-string*) (Olins & Olins, 1974). Esta estructura primaria de la cromatina (fibra de 10 nm) puede compactarse en agregados o fibras de orden superior gracias a la interacción entre histonas de diferentes nucleosomas y/o a la presencia de proteínas *linker*, como la histona H1, la cual interacciona con los segmentos de ADN de entrada y salida del nucleosoma (Simpson, 1978).

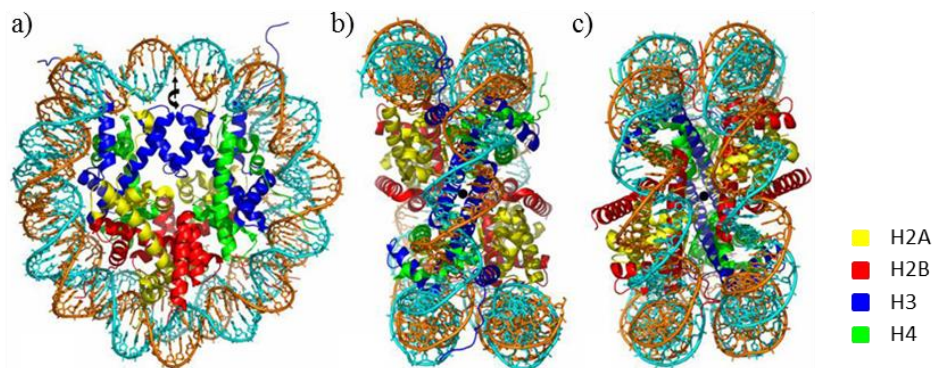


Figura I19. Estructura cristalográfica del nucleosoma canónico. 147 pb de ADN completan cerca de dos vueltas alrededor de un octámero de histonas H3, H4, H2A y H2B (a) Representación lateral, (b) superior y (c) inferior (Davey *et al.*, 2002).

Box3: Histona H1

Las histonas *linker* o H1 interaccionan con los fragmentos de ADN de entrada y salida del nucleosoma. Esta interacción da estabilidad al nucleosoma y a las conformaciones de empaquetamiento superiores de la cromatina (Figura I20). Hay distintos subtipos de histona H1 y se han evidenciado funciones específicas para alguno de ellos. Por ejemplo, la histona H1 de línea germinal juega un papel importante durante los primeros estadios del desarrollo embrionario y la reprogramación génica (revisado en Fyodorov *et al.*, 2018).

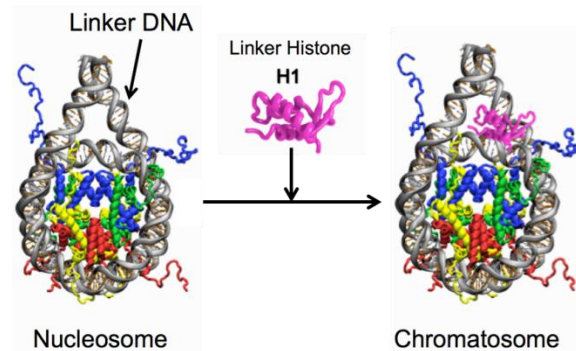


Figura I20. Modelo de unión de la histona *linker* H1. La histona H1 o histona *linker* se une a los segmentos de entrada y salida del ADN dando estabilidad al nucleosoma (Adaptación de Draizen *et al.*, 2016).

El plegamiento de la fibra de nucleosomas en estructuras de órdenes superiores aún se encuentra en amplio debate. La microscopía electrónica (EM) permitió observar que algunas fibras de 10 nm se repliegan sobre sí mismas, formando fibras de 30 nm de diámetro. La regularidad de estas fibras dependía de la secuencia del ADN, la longitud del ADN *linker* y las condiciones experimentales. Estos análisis llevaron a proponer dos modelos de empaquetamiento. El primero es una estructura solenoide donde la fibra de 10 nm forma una hélice con 6-8 nucleosomas por vuelta (Robinson *et al.*, 2006). En el segundo modelo, la fibra sigue una disposición en zigzag entre los nucleosomas adyacentes (Dorigo *et al.*, 2004). Sin embargo, nuevas técnicas y observaciones con microscopía *in vivo* han revelado que la fibra de 10 nm se repliega generalmente de manera irregular formando “*clutches*”, es decir, grupos heterogéneos de nucleosomas de tamaño y densidad variable (Ricci *et al.*, 2015) (Figura I21).

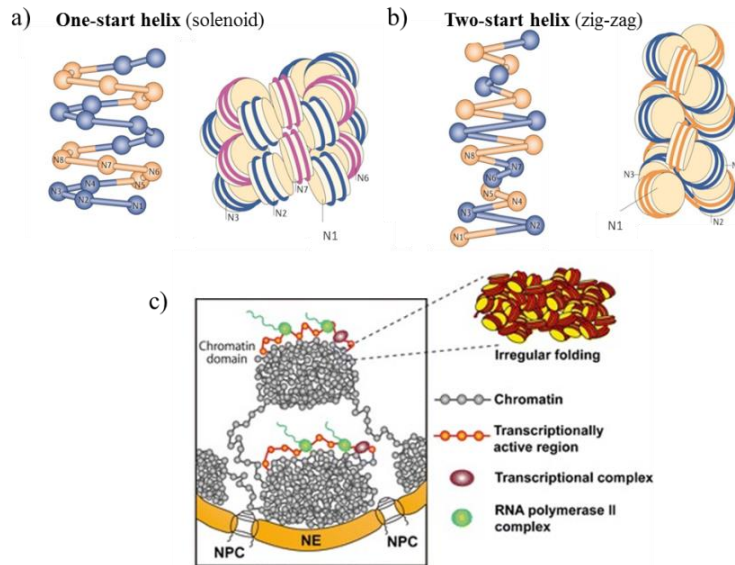


Figura I21. Modelos de compactación de la cromatina. (a) y (b) Modelos clásicos de plegamiento de la fibra de nucleosomas siguiendo una estructura regular (Luger *et al.*, 2012). (c) Modelo irregular de la cromatina formando “clutches” (NE = membrana nuclear, NPC = complejo del poro nuclear) (Maeshima *et al.*, 2010).

2.2 Variantes de Histonas

Aun siendo los nucleosomas un complejo altamente conservado evolutivamente, la presencia de variantes de histonas y sus múltiples modificaciones post-traduccionales, hacen de los nucleosomas unas entidades únicas, con inestabilidad posicional y fluctuaciones conformacionales específicas que condicionan la estructura y función local de la cromatina (Zentner & Henikoff, 2013; Lai & Pugh, 2017). Mientras las histonas canónicas se encuentran en todos los organismos eucariotas y su producción está ligada a la replicación del ADN, las variantes de histonas se expresan de forma constitutiva y son variedades específicas de linaje evolutivo (Talbert & Henikoff, 2010). Así, por ejemplo, en *S. cerevisiae*, sólo hay dos variantes descritas: Htz1 y Cse4, variantes de las histonas H2A y H3, respectivamente. En cambio, en mamíferos, se han descrito más de 20. Las variantes de histonas se encuentran asociadas a un amplio abanico de procesos celulares como la reparación, recombinación meiótica, segregación cromosómica, inicio y fin de la transcripción, condensación de los cromosomas sexuales y el empaquetamiento de la cromatina espermática (Figura I22).

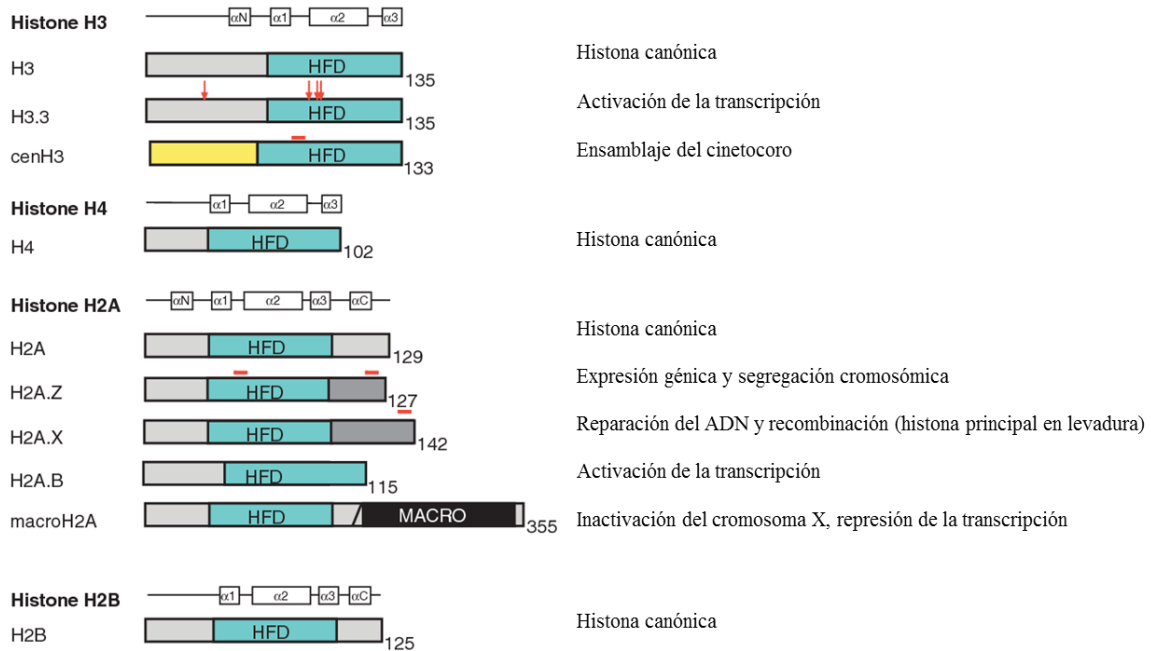


Figura I22. Variantes de histonas. Principales variantes de las histonas H3, H4, H2A y H2B junto a sus funciones asociadas. Los cambios en la secuencia entre variantes se indican en rojo (Adaptación de Henikoff & Smith, 2015).

Box4: Cse4: Variante histona H3.

Una característica común de los centrómeros eucariotas es la presencia de nucleosomas en los que la histona canónica H3 es substituida por la variante cenH3 (CENP-A en humanos y Cse4 en *S. cerevisiae*) (Malik & Henikoff, 2009). La arquitectura adoptada por los nucleosomas centroméricos permite el ensamblaje del cinetocoro y la unión de los microtúbulos del huso mitótico, por lo que su papel es esencial en la segregación cromosómica durante la división celular (Malik & Henikoff, 2003). Las histonas cenH3 presentan un gran nivel de divergencia entre las especies eucariotas y se caracterizan por colas N-terminal de longitudes variables. Aun así, la variante Cse4 es capaz de reemplazar funcional y estructuralmente CENP-A en células humanas (Wieland et al., 2004).

2.3 Modificaciones postraduccionales

Tanto las histonas canónicas como las variantes presentan modificaciones postraduccionales (PTMs), siendo las acetilaciones, metilaciones y fosforilaciones las más descritas. Las PTMs tienen lugar, principalmente, en los extremos C- y N- terminal debido

Hexasoma	Partícula con un tetrámero central 2(H3-H4) junto con una sola copia del dímero H2A-H2B.
Tetrasoma	Partícula formada exclusivamente por el tetrámero central 2(H3-H4).
Hemisoma	Partícula formada por una copia de cada una de las histonas H4, H2A, H2B y CenH3, variante de H3 específica de centrómero, formando un tetrámero heterotípico.
R-Octasoma	Octasoma canónico de histonas pero con el ADN envuelto en sentido dextrógiro (reversoma).
R-Tetrasoma	Partícula formada por un tetrámero central 2(H3-H4) con el ADN en sentido dextrógiro
Partículas NHP	Partículas que contienen proteínas no-histona (<i>non-histone proteins</i>) en sustitución de H2A y H2B

Tabla I1. Nucleosomas no canónicos descritos. Tabla resumen de los nucleosomas no canónicos descritos en la cromatina.

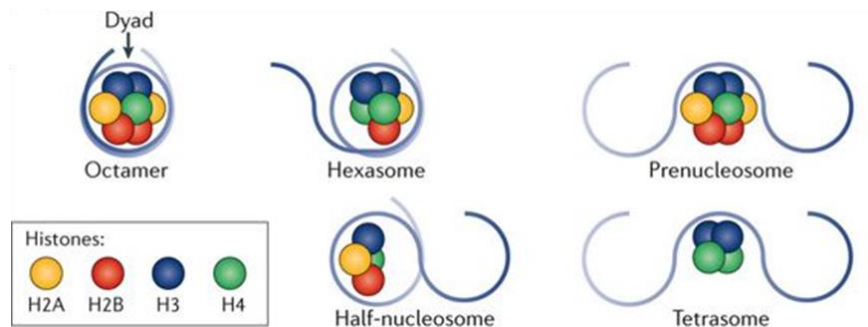


Figura I24. Nucleosomas no canónicos. Esquema representativo de algunas de las partículas “nucleosome-like” descritas con variación en número o composición de las histonas (Lai & Pugh, 2017)

2.5 Posición de los nucleosomas

A partir de estudios realizados con nucleasa micrococcal (MNasa), ADNasa I y técnicas de alta resolución (secuenciación masiva y ChiP-seq) sobre una gran variedad de organismos, desde levadura a células humanas, se ha mapeado la posición de los nucleosomas, junto con sus variantes de histonas y las PTMs que éstas puedan presentar (Yuan *et al.*, 2005; Mavrich *et al.*, 2008a; Valouev *et al.*, 2008; Rodriguez *et al.*, 2014). *In vivo*, la posición de los nucleosomas parece estar determinada por varios factores, como son la secuencia del ADN, la presencia de otras proteínas de unión al ADN, y la actividad de complejos remodeladores de la cromatina. Recíprocamente, la presencia de nucleosomas puede bloquear el acceso a las ARN-polimerasas y la unión al ADN de otros complejos y factores reguladores.

En el caso de *S. cerevisiae*, organismo modelo en esta tesis, hay aproximadamente 60.000 nucleosomas, los cuales pueden variar ligeramente su posición (hasta ± 50 pb) e incluso superponer su localización (Fragoso *et al.*, 1995; Tanaka *et al.*, 1996). Comparando todos los genes de *S. cerevisiae*, los nucleosomas siguen ciertos patrones de posicionamiento respecto al lugar de inicio de la transcripción (TSS) (Mavrich *et al.*, 2008b). Junto al extremo 5' del TSS se encuentra, por lo general, una región de entre 50 y 150 pb aparentemente libre de nucleosomas (NFR, *nucleosome free region*) (Sekinger *et al.*, 2005). La NFR está flanqueada, en el lado 3', por un nucleosoma estable y bien posicionado (nucleosoma +1), que ya corresponde a la región codificante del gen. Por el lado 5' (promotor), se pueden posicionar uno o dos nucleosomas menos estables (-1 y -2). Esta región del promotor contiene secuencias reguladoras (p. e. sitios de unión a factores de transcripción) por lo que la presencia o no de estos nucleosomas (-1 y -2) junto con sus modificaciones influirán sobre la actividad transcripcional del gen (Yuan *et al.*, 2005; Barski *et al.*, 2007). En el resto del gen, los nucleosomas de la región transcrita se disponen a partir del nucleosoma +1, adoptando posiciones más difusas (*fuzzy nucleosomes*) a medida que nos alejamos del TSS. Finalmente, en el extremo 3' de la región transcrita volvemos a encontrar una región libre de nucleosomas (NFRII), donde generalmente finaliza la transcripción por parte de la ARN-polimerasa II (TTS). Esta región suele estar precedida por un nucleosoma bien posicionado (Figura I25a). En los orígenes de replicación de *S. cerevisiae*, también se observan patrones de posicionamiento característicos, con una región NFR en las secuencias ARS (secuencia de replicación autónoma) y nucleosomas bien posicionados en ambos flancos (Figura I25b).

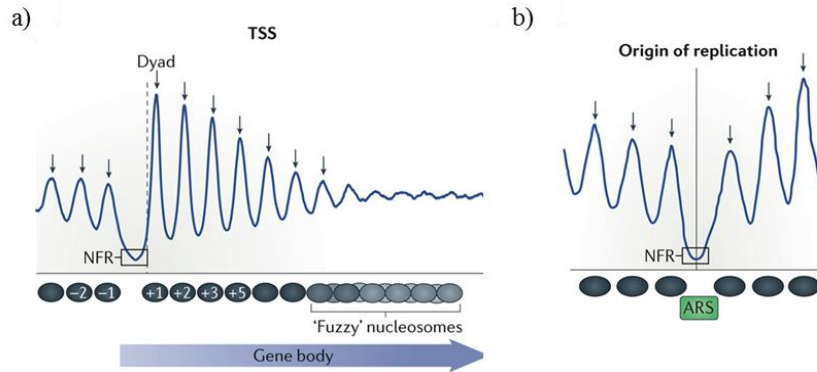


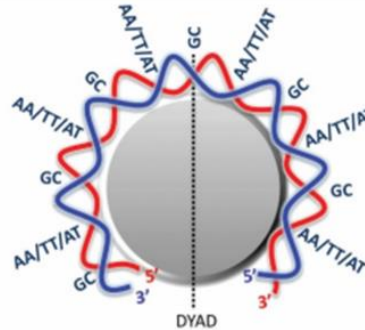
Figura I25. Patrón de posicionamiento de los nucleosomas en *S. cerevisiae*. (a) Esquema del patrón nucleosomal observado en los genes de *S. cerevisiae* respecto al lugar de inicio de la transcripción (TSS) con la región libre de nucleosomas (NFR) flanqueada por los nucleosomas -2, -1 y +1. A medida que nos adentramos en la región transcrita, el posicionamiento de los nucleosomas se vuelve más difuso. (b) Esquema del patrón nucleosomal del origen de replicación en levadura. La secuencia ARS corresponde a una región NFR flanqueada por nucleosomas bien posicionados. Los picos y valles en los gráficos indican las ubicaciones de alta y baja ocupación de los nucleosomas, respectivamente (Lai & Pugh, 2017).

2.6 Determinantes del posicionamiento nucleosomal

Los estudios de posicionamiento de nucleosomas han permitido determinar patrones a lo largo de la secuencia del ADN y caracterizar aquellas que facilitan o excluyen la formación de nucleosomas. Por ejemplo, las secuencias poli-dA/dT, que forman una doble hélice de ADN más rígida con un surco menor más pequeño, dificultan la acomodación del nucleosoma (Segal & Wisdom, 2009; García *et al.*, 2017). Esta característica se encuentra en el 70% de las NFRs de promotores en *S. cerevisiae*, lo que permite regiones de ADN libres para la unión de los factores de transcripción (Iyer & Struhl, 1995). De forma contraria, las regiones con un patrón periódico de 10 pb en la distribución de dinucleótidos AA/TT/AT son más flexibles para curvarse, por lo que tienen una mayor probabilidad de posicionamiento de nucleosomas (Figura I26). En los segmentos de entrada y salida de los nucleosomas, se observa también un mayor contenido en adeninas y timinas en los extremos 5' y 3', respectivamente. Esta distribución asimétrica de los nucleótidos y secuencias más y menos flexibles, implica un código estructural en el ADN que determina el patrón de posicionamiento nucleosomal y

la organización de la cromatina de forma especie-específica (Segal *et al.*, 2006; Washietl *et al.*, 2008; García *et al.*, 2017).

Figura I26. Flexibilidad del ADN y posicionamiento de los nucleosomas. Las secuencias de ADN con motivos AA/TT/AT alternados con dinucleótidos GC permiten una mejor acomodación de los nucleosomas debido a la curvatura periódica del surco menor de la doble hélice



Además de la secuencia, los complejos remodeladores de cromatina, los factores de transcripción y las RNA polimerasas modifican el posicionamiento de los nucleosomas (Yarragudi *et al.*, 2004; Mavrich *et al.*, 2008b; Rando & Winston, 2012). Los remodeladores son proteínas tipo chaperona que se unen a las histonas evitando su agregación (Hammond *et al.*, 2017), facilitan su deslizamiento (*sliding*) y su recambio (*replacement*) (Partensky & Narlikar, 2009) (Figura I27). Los factores de inicio de la transcripción y de la replicación (ORC) pueden actuar como una barrera que fijaría la posición de los nucleosomas colindantes.

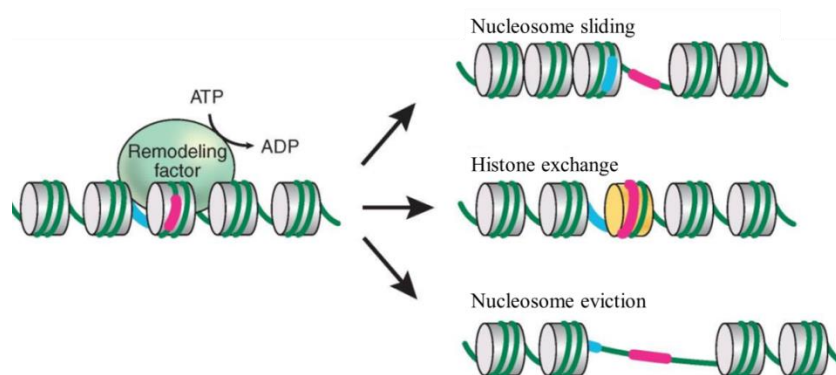


Figura I27. Complejos remodeladores de la cromatina. Representación de algunos de los cambios sobre los nucleosomas (*sliding*, recambio o eliminación de histonas) llevados a cabo por los complejos remodeladores de la cromatina (Becker & Workman, 2013).

2.7 Dinámica conformacional de los nucleosomas

A parte de las transiciones llevadas a cabo por complejos remodeladores de la cromatina como el deslizamiento o reemplazamiento de los nucleosomas (*sliding, replacement*), los nucleosomas presentan internamente una notable dinámica conformacional. Estudios *in vitro* usando nucleasas y, posteriormente, la transferencia de energía por resonancia fluorescente (FRET), indican cambios rápidos en la interacción de los segmentos de ADN de entrada y salida del nucleosoma (*breathing*). Estos cambios pueden dar lugar a un desenrollamiento parcial (*unwrapping*) del ADN sobre el octámero de histonas y hacer accesibles secuencias internas del ADN (Li & Widom, 2004; Van Holde & Zlatanova, 2006). Asimismo, los dímeros H2A/H2B y el tetrámero H3/H4 pueden disociarse parcialmente debido a una unión más débil entre ellos (Baxevanis *et al.*, 1991). En ese caso, se produce el desenrollamiento parcial del ADN alrededor del nucleosoma (*opening o unfolding*), un cambio conformacional importante durante la transcripción (Kireeva *et al.*, 2002) (Figura I28). Estas fluctuaciones convierten a los nucleosomas en unas unidades estructurales muy dinámicas. Las variantes de histonas y las PTMs, además, pueden influir marcadamente en el abanico de conformaciones posibles. Por tanto, además de la posición, la dinámica conformacional de los nucleosomas condiciona la estructura y actividades a lo largo de la cromatina (Rando & Ahmad, 2007).

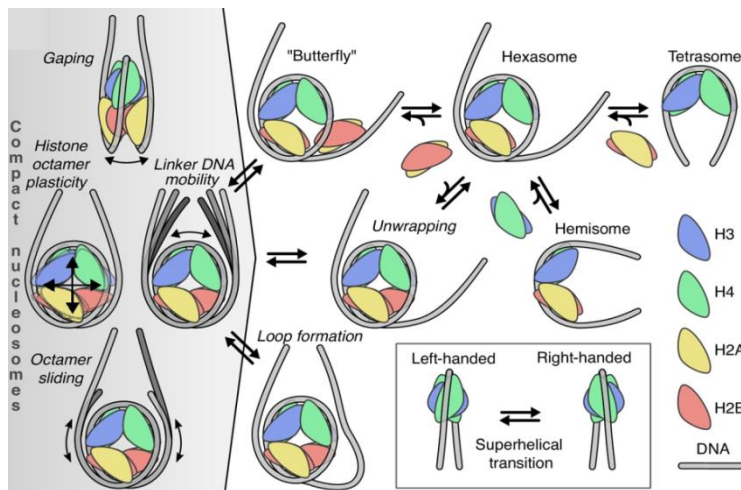


Figura I28. Dinámica conformacional de los nucleosomas. Esquema representativo de las conformaciones adoptadas por los nucleosomas según el grado y dirección del enrollamiento del ADN o la disociación parcial de las histonas (Armeev *et al.*, 2019).

3. TOPOLOGÍA DEL ADN NUCLEOSOMAL

Mientras las características bioquímicas y propiedades biofísicas de los nucleosomas han sido ampliamente estudiadas, sus implicaciones en la topología del ADN no han sido del todo descifradas. Sin embargo, el enrollamiento del ADN sobre el octámero de histonas es un factor importante en la estabilización, amortiguación y modulación de la superhelicidad del ADN intracelular.

No fue hasta una década después de la observación del superenrollamiento de las moléculas circulares de ADN del poliomavirus SV40 (Vinograd *et al.*, 1965), que se entendió que este superenrollamiento era debido a la presencia de nucleosomas. La extracción de las histonas de los minicromosomas de SV40 en la célula huésped era lo que generaba moléculas de ADN superenrolladas (Germond *et al.*, 1975). Mediante el estudio comparativo de las distribuciones electroforéticas de Lk de moléculas de ADN circulares de SV40 con y sin nucleosomas se calculó un valor de $\Delta Lk = -26$ unidades (Keller, 1975; Shure & Vinograd 1976). Este valor obtenido resultó ser, aproximadamente, igual al número promedio de nucleosomas observados mediante EM (Saragosti *et al.*, 1980; Sogo *et al.*, 1986). Con estos resultados se concluyó que, en los minicromosomas de SV40, el valor de Lk disminuía aproximadamente una unidad por nucleosoma ($\Delta Lk = -1$) (Figura I29). Experimentos similares con cromatina nativa de minicromosomas circulares de *S. cerevisiae* obtuvieron el mismo valor de ΔLk por nucleosoma. Por ejemplo, en el minicromosoma circular TRP1ARS1 se posicionan 7 nucleosomas y su ΔLk se determinó cercano a -7 (Pederson *et al.*, 1986). Finalmente, experimentos *in vitro* de reconstrucción de cromatina, utilizando las 4 histonas del octámero en secuencias específicas de ADN dispuestas en tándem, obtuvieron también valores similares, $\Delta Lk = -1.01 \pm 0.08$ (Simpson *et al.*, 1985) y $\Delta Lk = -1.04 \pm 0.08$ (Norton *et al.*, 1989) por nucleosoma.

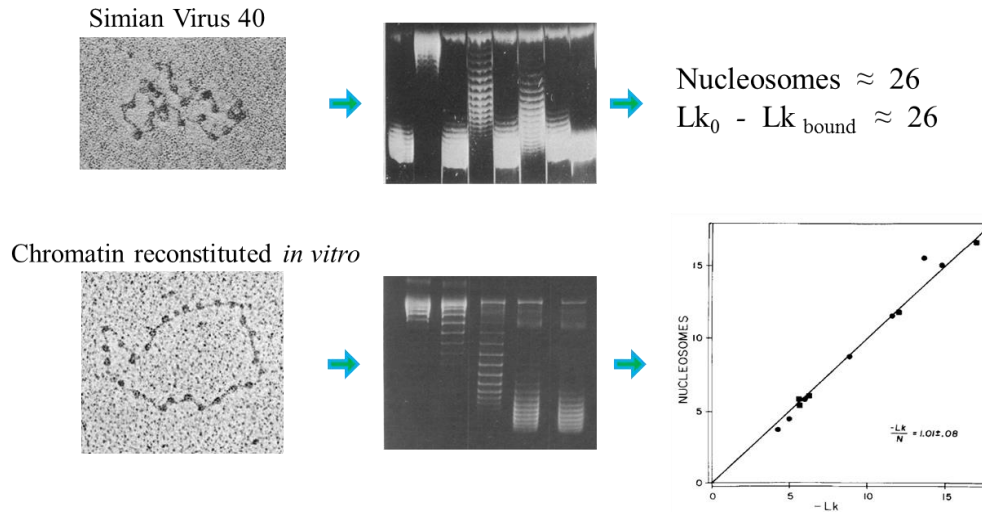


Figura 129. Primeras aproximaciones del valor de ΔLk nucleosomal. Al dividir el valor de Lk obtenido en los análisis electroforéticos por el número de nucleosomas observados mediante microscopia electrónica se obtuvo el valor paradójico de $\Delta Lk = -1/\text{nucleosoma}$.

Box5: Minicromosoma TRP1ARS1

El minicromosoma TRP1ARS1 es un pequeño minicromosoma circular utilizado como modelo de estudio de la cromatina. Se genera *in vitro* a partir de la circularización de un segmento genómico de *S. cerevisiae* de 1453 pb del cromosomas IV, situado entre dos dianas de restricción EcoRI. En este segmento se encuentra el gen TRP1 junto a la secuencia autónoma de replicación ARS1, lo que permite su transfección y mantenimiento en cepas $\Delta trp1$ en forma de anillo extracromosómico. El minicromosoma TRP1ARS1 mantiene la misma organización nucleosomal presente en el cromosoma IV (Figura 130).

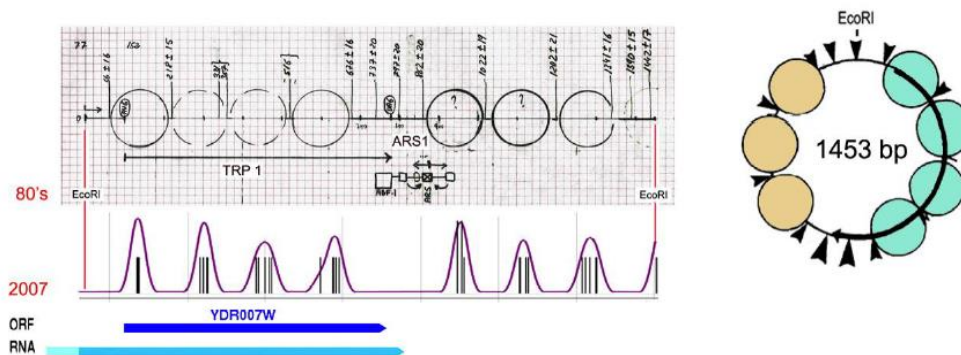


Figura 130. Minicromosoma TRP1ARS1 de *S. cerevisiae*. Mapa del posicionamiento nucleosomal de la región TRP1ARS1 en forma de minicromosoma circular (arriba) y en su ubicación genómica (abajo) junto con la representación esquemática del anillo con los nucleosomas y las regiones susceptibles de ser digeridas por la MNasa (flechas).

3.1 La paradoja del número de enlace

Los estudios biofísicos y bioquímicos del nucleosoma (Noll, 1974; Travers & Klug, 1987; Zivanovic *et al.*, 1988), así como las estructuras cristalográficas obtenidas mediante difracción con rayos X (Finch *et al.*, 1977; Richmond *et al.*, 1984), demostraron que el ADN describe cerca de dos vueltas levóginas alrededor del octámero de histonas. Esta geometría implicaba un valor de ΔWr cercano a -2 para el ADN nucleosomal. Mediante estos estudios se concluyó también que los cambios en el número de torsión helicoidal del ADN nucleosomal eran poco significativos ($\Delta Tw \approx 0$). Con estos valores y teniendo en cuenta la ecuación general del Lk ($\Delta Lk = \Delta Tw + \Delta Wr$) se esperaría un valor de ΔLk cercano a -2 unidades por nucleosoma. Sin embargo, las mediciones experimentales del número de enlace estabilizado por los nucleosomas ($\Delta Lk = -1$) diferían marcadamente del valor inferido mediante los cálculos de $\Delta Wr + \Delta Tw = \Delta Lk \approx -2$. Esta discrepancia dio lugar a la llamada "Paradoja del número de enlace" o "*Linking number paradox*" que fue centro de debate durante los años 80 y 90 del siglo pasado.

3.2 Hipótesis para la resolución de la paradoja del número de enlace

La falta de correspondencia entre los valores obtenidos de forma experimental ($\Delta Lk \approx -1$) y el valor teórico ($\Delta Lk \approx -2$) dio lugar, en su momento, a dos hipótesis no excluyentes como posible explicación de la paradoja del número de enlace.

3.2.1 Overtwisting del ADN nucleosomal

La primera hipótesis para explicar la paradoja postuló que, mientras que el número de enrollamiento ΔWr era cercano a -1.8, el número de torsión del ADN del *core* nucleosomal era notable y positivo ($\Delta Tw \approx +0.8$). Por tanto, $\Delta Wr + \Delta Tw = \Delta Lk \approx -1$. El aumento de ΔTw vendría dado por la disminución del *helical repeat* del ADN del *core* ($h \leq 10.0$ pb/vuelta) en comparación al valor de referencia del ADN desnudo en disolución ($h \approx$

10.5 pb/vuelta). Este *overtwisting* podría darse como consecuencia de la interacción directa del ADN con las histonas, las cuales estabilizarían esta deformación de la doble hélice (Hayes *et al.*, 1991).

Para comprobar si realmente $\Delta Tw \approx +0.8$, se calculó con precisión este valor mediante la ecuación:

$$\Delta Tw = \Delta \Phi + \Delta STw$$

El denominado *winding number* (Φ) describe la torsión de la hélice de ADN alrededor de su eje a partir de una superficie de referencia y depende de h_s (*surface-related helical repeat*). Mediante estudios experimentales de *footprinting* con ADNasa I, escisión de radicales hidroxilo y de periodicidad de pares de bases se estimó un valor promedio de $h_s \approx 10.2$ pb/vuelta (Prunell *et al.*, 1979; Hayes *et al.*, 1991). Tomando este valor y la longitud del ADN del *core* nucleosomal (146 pb), se obtuvo el valor de $\Phi = 14.3, +0.4$ unidades respecto a un ADN lineal de 146 pb. El denominado *surface twist* (*STw*) es una corrección para tener en cuenta la curvatura del ADN alrededor del octámero. Su valor es -0.19 (White & Bauer, 1989). Con estos datos, la ecuación $\Delta \Phi + \Delta STw$ indicaba un ΔTw de +0.2 para el ADN del nucleosoma, lejos del valor +0.8 esperado.

Posteriormente, la resolución atómica de la estructura del nucleosoma (Luger *et al.*, 1997) permitió medir ΔTw con precisión y se corroboró el valor de +0.2. Además, se observó que el ADN realiza 1.65 vueltas levóginas alrededor del octámero de histonas con un ángulo de paso de 4° . Estos valores permitieron calcular también de forma precisa el ΔWr del ADN nucleosomal, el cual se estableció, ahora, en -1.5 unidades (Prunell, 1998). Con estos nuevos valores el ΔLk esperado sería: $0.2 - 1.5 \approx -1.3$ unidades, un valor más cercano a -1 pero aún insuficiente para explicar la paradoja del *Lk*.

3.2.2 Contribución del ADN linker

La segunda hipótesis para resolver la paradoja del número de enlace se basa en la topología de los segmentos de ADN *linker* y su orientación entre nucleosomas contiguos. En la organización observada en forma de zigzag, los nucleosomas se apilan los unos sobre los otros y los segmentos de ADN de entrada y salida del nucleosoma se cruzan continuamente siguiendo un mismo patrón. Se postuló que esta geometría reiterada aumentaría el valor global de ΔWr de la fibra en zigzag dando lugar al valor de $\Delta Lk = -1$ por nucleosoma (Worcel *et al.*, 1981). Sin embargo, estudios recientes de modelado de fibras en zigzag y experimentos *in vitro* de reconstrucción de *arrays* regulares de nucleosomas, demuestran que el valor de ΔLk de estas fibras regulares varía marcadamente según la longitud del ADN *linker* (L). El valor de ΔLk variaba de -1.4 unidades en *arrays* con $L = 20$ pb a -0.8 unidades en $L = 25$ pb (Nikitina *et al.*, 2017) (Figura I31).

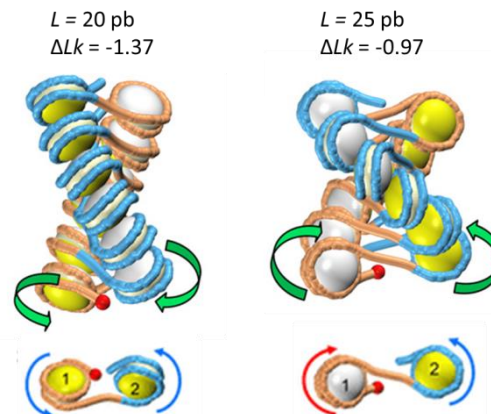


Figura I31. Contribución del ADN linker. Modelado de fibras en zigzag donde se muestra la variabilidad en el valor de Lk por nucleosoma en función de la longitud del ADN *linker* (L) (Nikitina *et al.*, 2017).

Relacionado con esta segunda hipótesis, estudios de nucleosomas individuales reconstituidos en pequeños anillos de ADN (De Lucia *et al.*, 1999) y estudios de la respuesta a la torsión de *arrays* de nucleosomas mediante pinzas magnéticas (Bancaud *et al.*, 2006) también apuntaron a la orientación del ADN fuera del *core* nucleosomal como explicación para la paradoja del Lk . Estos estudios sugirieron que los nucleosomas fluctúan entre tres conformaciones diferentes: una donde los segmentos de ADN de entrada y

salida del *core* se cruzan estableciendo un cruce negativo con $\Delta Wr = -1.8$ (estructura “canónica” del nucleosoma), una conformación abierta donde los segmentos no se cruzarían y el nucleosoma adoptaría una disposición parcialmente desenrollada ($\Delta Wr = -1.4$) y una tercera conformación donde los segmentos formarían un cruce positivo ($\Delta Wr \approx 0$) (Figura I32). Como resultado de la dinámica entre las diferentes conformaciones se obtendría un valor promedio de $\Delta Lk = -1$. El punto débil de esta hipótesis es que implica la necesidad de fuerzas externas significativas para variar la orientación de los segmentos de entrada y salida y, por tanto, la fluctuación entre las tres conformaciones.

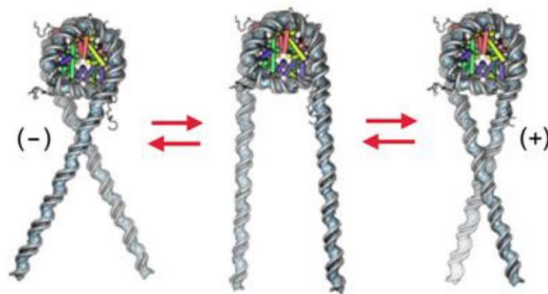


Figura I32. Fluctuación de los nucleosomas. Representación de las tres conformaciones que pueden adoptar los nucleosomas en función de la posición de los segmentos de ADN de entrada y salida. De izquierda a derecha: cruce negativo (-) ($\Delta Wr = -1.8$), conformación abierta ($\Delta Wr = -1.4$) y cruce positivo (+) ($\Delta Wr \approx 0$).

3.3 Topologías de nucleosomas atípicos: el Centrómero puntual de *Saccharomyces cerevisiae*.

Los experimentos de topología citados hasta el momento sólo incluían o consideraban nucleosomas canónicos. Sin embargo, hay nucleosomas con componentes y propiedades específicas como es el caso de los nucleosomas de las regiones centroméricas. En particular, cada uno de los 16 cromosomas de *S. cerevisiae* posee un centrómero puntual que consiste en un único nucleosoma. La posición y estructura de este nucleosoma vienen determinados por la secuencia de un segmento de aproximadamente 200 pb de ADN, que incluye tres elementos: CDEI, CDEII y CDEIII (Bloom

& Carbon, 1982; Clarke, 1998). El elemento central, CDEII, es el que interacciona con el *core* de histonas, donde la variante de H3, CENP-A o *Cse4* en el caso de levadura, sustituye la histona H3. El elemento CDEI interacciona con el factor de transcripción Cbf1, que da estabilidad al centrómero puntual. Y, CDEIII interacciona con el complejo multiprotéico CBF3 (Ndc10, Cep3, Ctf13, Skp1), el cual interacciona con el cinetocoro (Xiao *et al.*, 2011; Cho & Harrison, 2011) (Figura I33).

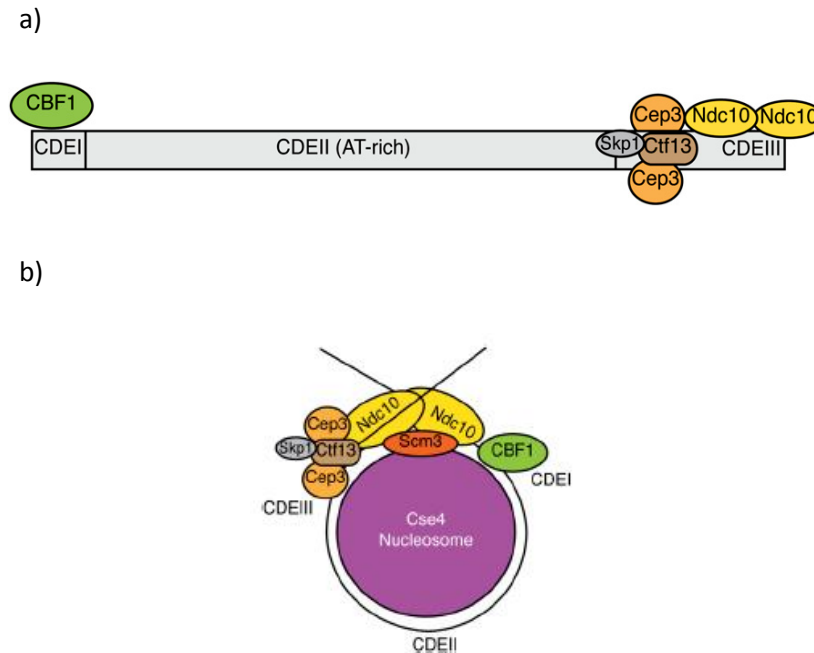


Figura I33. Esquema del centrómero puntual de *S. cerevisiae*. (a) Esquema de las secuencias determinantes para el posicionamiento del centrómero (CDEs). CDEI (8-10 pb): interacciona con la proteína Cbf1. CDEII (80-90 pb): región rica en AT que interacciona con el núcleo de histonas, en el que se incluye *Cse4* (variante centromérica de H3). CDEIII (26 pb): interacciona con el complejo CBF3 (Ndc10, Cep3, Ctf13, Skp1). (b) Modelo estructural del nucleosoma donde el homodímero Ndc10 interacciona con CDEIII y CDEI para formar un *loop* en CDEII. Ndc10 recluta una chaperona (Smc3) que deposita *Cse4* y otras histonas en CDEII.

La composición del centrómero puntual de *S. cerevisiae* sugiere una topología del ADN diferente a la de un nucleosoma canónico. En 1984, Bloom *et al.* analizaron la topología de minicromosomas circulares de *S. cerevisiae* con y sin centrómero. En los experimentos publicados se observó que los minicromosomas con centrómero presentaban un valor ΔLk menos negativo que los minicromosomas sin centrómero. Análisis similares sobre el

plásmido 2-micras de *S. cerevisiae*, el cual también posee un centrómero puntual (Huang *et al.*, 2011), mostraron de nuevo valores de ΔLk menos negativos de lo esperado bajo el supuesto de que los restantes nucleosomas estabilizan un valor ΔLk cercano a -1. Estas observaciones fueron corroboradas posteriormente por Furuyama & Henikoff en 2009, los cuales propusieron que el ADN del centrómero puntual de *S. cerevisiae* tendría una topología singular por estar enrollado en sentido dextrógiro ($\Delta Wr > +1$) en lugar de sentido levógiro. Como el segmento CDEII sólo abarca en torno a 80 pb, estos autores también postularon que el complejo de histonas asociado no sería un octasoma, sino que se trataría de un hemisoma formado por una copia de las histonas H2A, H2B, H4 y la variante CENP-A (CenH3).

4. HIPÓTESIS DE PARTIDA

La paradoja del número de enlace del ADN nucleosomal se originó a finales de los años 70 del siglo pasado cuando, al medir experimentalmente la diferencia de Lk estabilizada por nucleosoma, se encontró un valor ΔLk próximo a -1. Generalmente, para el cálculo de este valor, primero se determinaba, mediante microscopía electrónica, el número de nucleosomas presentes en ADN circulares, como son el cromosoma del virus SV40 o los minicromosomas de *S. cerevisiae*. Posteriormente, se comparaba, mediante electroforesis, la distribución de los topoisómeros del ADN circular relajado con topoisomerasa I en presencia y en ausencia de nucleosomas, con lo que se determinaba la diferencia de Lk entre las dos condiciones. Esta diferencia se dividía por el número de nucleosomas observados para obtener así el ΔLk medio estabilizado por nucleosoma.

En nuestro laboratorio, llevamos a cabo una revisión bibliográfica de los trabajos relacionados y se encontró que, tanto el conteo del número de nucleosomas como el cálculo de los valores de ΔLk , se trataban de aproximaciones y presentaban ciertos errores experimentales. Como ejemplo y de forma resumida:

- El número exacto de nucleosomas presentes en un mismo ADN circular (p.e. SV40) solía variar según el laboratorio, el origen de la muestra (cromatina nativa o *in vitro*), las condiciones de manipulación de la cromatina, o la técnica de preparación para microscopía.
- La relajación con topoisomerasa I de muestras de ADN, con y sin nucleosomas, no siempre se realizaba en las mismas condiciones y la misma temperatura. Por tanto, el valor h (*helical repeat*) de referencia variaba, lo que podía generar errores al calcular ΔLk .
- El análisis de las distribuciones de topoisómeros mediante electroforesis no siempre determinaba el valor medio de las distribuciones de Lk . En algunos

estudios se consideraba al ADN circular abierto (con *nicks*) como la forma relajada, lo que es incorrecto para el cálculo de ΔLk .

Por otro lado, en estos estudios se asumía que todas las partículas visibles en la fibra de 10 nm correspondían a nucleosomas canónicos (formados por el octámero típico de histonas). Sin embargo, a día de hoy sabemos que la cromatina nativa puede presentar nucleosomas especiales, como el caso del centrómero, y otros complejos multiprotéicos que pueden confundirse con nucleosomas. En el caso de las fibras de 10 nm reconstituidas *in vitro*, también pueden presentar complejos sub-octaméricos, cuya topología del ADN difiere de los nucleosomas octaméricos.

Todas estas observaciones llevaron a nuestro laboratorio a la conclusión de que el valor generalmente asumido, $\Delta Lk = -1$ por nucleosoma, no era en absoluto preciso. Además, los numerosos análisis topológicos de diversos minicromosomas circulares de *S. cerevisiae* realizados en nuestro laboratorio nos permitieron corroborar que:

- Los minicromosomas presentaban, tanto *in vitro* como *in vivo*, valores ΔLk más negativos de lo que cabría esperar si se considera que $\Delta Lk = -1$ por nucleosoma.
- Independientemente de lo anterior, la presencia de un centrómero puntual producía valores de ΔLk menos negativos.

Las premisas anteriores nos llevaron a la **hipótesis de partida** para la realización de esta tesis: La llamada “paradoja del número de enlace del ADN nucleosomal” podría deberse al acúmulo de pequeños errores experimentales a la hora de calcular el ΔLk estabilizado por nucleosoma. Por lo que el valor podría ser más negativo que -1 para los nucleosomas canónicos, resolviendo la histórica contradicción. Mientras que en el nucleosoma centromérico este valor sería positivo.

5. ESTRATEGIA EXPERIMENTAL

Para determinar el valor de ΔLk estabilizado por un elemento que interacciona con el ADN (p.ej. un nucleosoma) se utiliza un ADN circular covalentemente cerrado. Para ello, se compararan las distribuciones de Lk obtenidas tras relajar esa molécula de ADN con una topoisomerasa tipo IB antes y después de la unión de ese elemento (Figura I34). Este principio se basa en que la interacción del elemento deformará el ADN en forma de ΔTw y ΔWr . Al tratarse de un ADN covalentemente cerrado, su Lk no variará. Por tanto, las deformaciones de ΔTw y ΔWr estabilizadas mediante la unión del elemento se acompañarán siempre de deformaciones compensatorias (y de signo contrario) más allá del lugar de interacción. Como las deformaciones compensatorias de ΔTw y ΔWr no estarán estabilizadas, estas podrán ser relajadas por la topoisomerasa I. Al eliminar el elemento unido, el ADN presentará las deformaciones ($\Delta Tw + \Delta Wr$) que estaban estabilizadas por el elemento (no relajables), las cuales producirán un desplazamiento de la distribución de valores de Lk (ΔLk). Este desplazamiento corresponderá a la "huella topológica" del elemento unido al ADN.

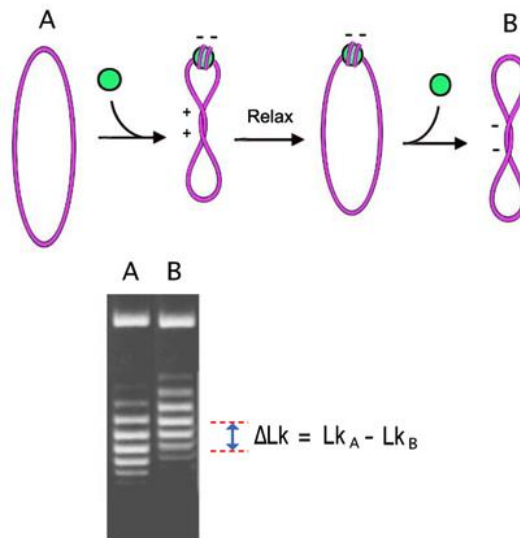


Figura I34. Principio de la "huella topológica". Ejemplo de como la unión de un elemento con el ADN relajado (A) estabiliza deformaciones $\Delta Tw + \Delta Wr = -2$, las cuales son compensadas mediante deformaciones $\Delta Tw + \Delta Wr = +2$, que pueden ser relajadas por una topoisomerasa I (no estabilizadas). Tras la disociación del elemento, el ADN presenta un $\Delta Lk = -2$ (B) como "huella topológica" de las deformaciones estabilizadas por el elemento. Los geles de electroforesis (inferior) permiten medir con precisión las diferencias en las distribuciones de Lk del ADN en las condiciones A y B, y por lo tanto, conocer el efecto del elemento unido.

Para aplicar el anterior principio, lo ideal es minimizar el tamaño de la molécula de ADN circular. Esto reduce la fluctuación conformacional del anillo, por lo que se generan distribuciones de Lk más estrechas (con menos topoisómeros). Con estas distribuciones, las diferencias de Lk correspondientes a distintas condiciones experimentales pueden calcularse con una mayor precisión. Por tanto, para analizar la topología de complejos presentes en cromatina nativa, lo ideal es utilizar minicromosomas circulares lo más pequeños posible. Este es el caso del minicromosoma TRP1ARS1 de *S. cerevisiae*, el cual contiene sólo 7 nucleosomas. Sin embargo, aún debería ser posible eliminar 2 de esos nucleosomas sin afectar el gen *TRP1* y el origen de replicación *ARS1*, obteniendo así un anillo autoreplicativo con 5 nucleosomas bien posicionados (Figura I35). Además, al reducirse el número de elementos cromatínicos se minimiza también el número de sus posibles variantes estructurales que alterarían el valor final de ΔLk por nucleosoma.

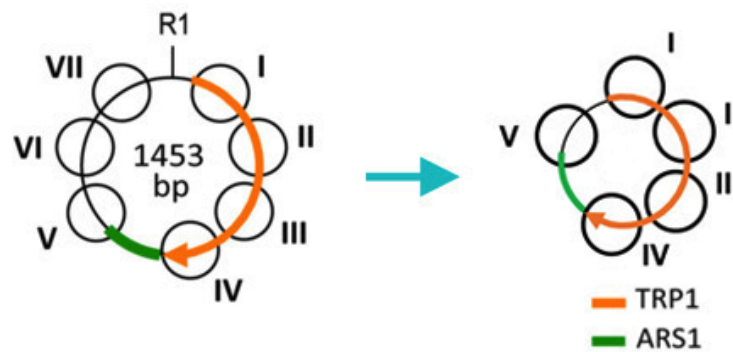


Figura I35. Organización de los nucleosomas en los minicromosomas. Esquema de la organización de los 7 nucleosomas (I- VII) presentes en el minicromosoma *TRP1ARS1* de *S. cerevisiae*. Tras eliminar la posición de los nucleosomas VI y VII, se generó un anillo autoreplicativo de sólo 5 nucleosomas (I-V) dónde se mantenían funcionales el gen *TRP1* y el origen de replicación *ARS1*.

Un segundo aspecto de la estrategia experimental propuesta consiste en optimizar las electroforesis de ADN y su posterior análisis para determinar con precisión el centro de cada distribución de valores de Lk (Figura I36). Mediante el uso de intercalantes (p. e.

cloroquina) ha de ser posible resolver y comparar en un mismo gel todos los topoisómeros de Lk generados en distintas condiciones experimentales. Si es necesario, la utilización de electroforesis bidimensionales permitirá evitar solapamientos de topoisómeros de Lk (covalentemente cerrados) con la forma circular abierta (*nicks*) y bandas de ADN linealizado. También permitirá enumerar correctamente las bandas correspondientes a los topoisómeros de Lk . Para determinar el centro de las distribuciones de valores de Lk , se medirá la intensidad (abundancia) de cada topoisómero por separado y se determinará entonces el centro de intensidad. La diferencia de Lk entre dos distribuciones será la distancia (en la escala Lk) entre los centros de ambas distribuciones (Figura I36).

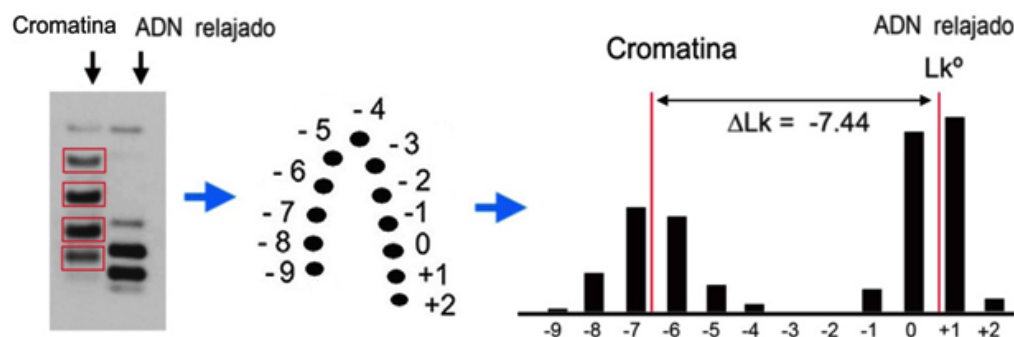


Figura I36. Determinación de ΔLk entre dos distribuciones. En el gel de la izquierda, el ajuste de las concentraciones de cloroquina permite la resolución de todos los topoisómeros de Lk de un minicromosoma (cromatina) junto con la distribución de Lk del mismo ADN relajado. Mediante electroforesis bidimensionales (esquema central) se puede enumerar sin ambigüedad la posición (el valor relativo) de los distintos topoisómeros de Lk al evitar solapamientos. Los histogramas de intensidades de cada topoisómero (derecha) permiten determinar el centro de cada distribución y, posteriormente, calcular la distancia entre los centros (ΔLk).

Dado un minicromosoma determinado de *S. cerevisiae*, el ΔLk estabilizado por los elementos de su cromatina nativa se obtendrá al comparar la distribución de Lk del minicromosoma fijado *in vivo* con la distribución de Lk del anillo de ADN desnudo y relajado *in vitro*. Para ello, es importante que tanto el minicromosoma *in vivo* como el

ADN relajado *in vitro* se encuentren a la misma temperatura en el momento de fijar su topología.

Una vez conocido el ΔLk estabilizado por la cromatina del minicromosoma, la pregunta es cómo determinar el ΔLk estabilizado por un elemento individual (un nucleosoma, un centrómero u otro complejo unido al ADN). La estrategia a seguir en este caso es insertar ese elemento i en el minicromosoma y calcular de nuevo el ΔLk estabilizado por la cromatina del minicromosoma + i . El ΔLk estabilizado por el elemento i será entonces la diferencia entre el ΔLk del minicromosoma antes y después de insertar el elemento i . Es decir, $\Delta\Delta Lk^i = \Delta Lk^{CRM+i} - \Delta Lk^{CRM}$ (Figura I37).

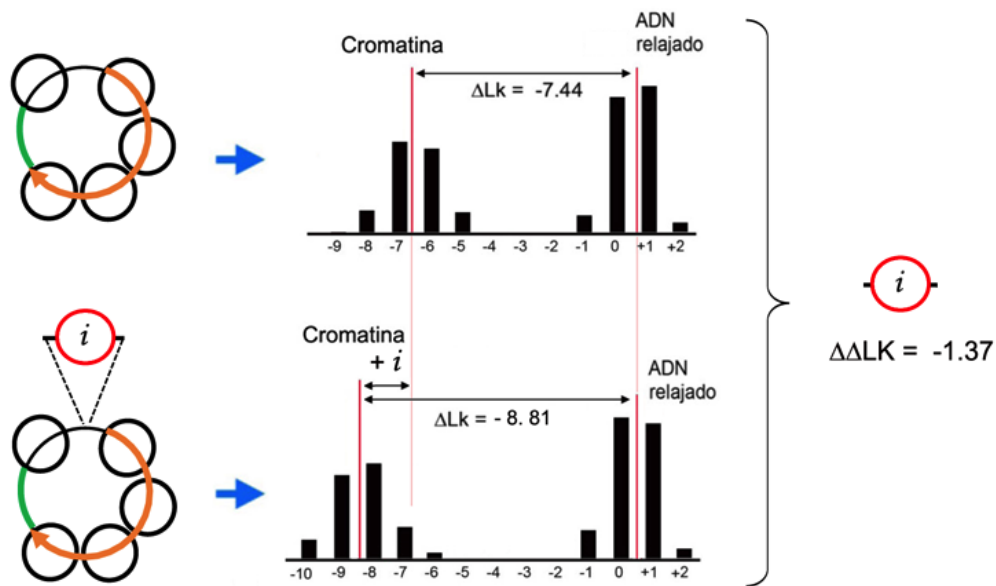


Figura I37. Determinación del ΔLk estabilizado por un elemento individual. La diferencia entre el valor de ΔLk de los minicromosomas con y sin el elemento (i) corresponde al ΔLk estabilizado por i . $\Delta\Delta Lk^i = \Delta Lk^{CRM+i} - \Delta Lk^{CRM}$

Dado que el minicromosoma circular es pequeño, el sitio y el espacio de inserción del elemento i son críticos, ya que no deben interferir con la estabilidad del minicromosoma ni alterar la posición del resto de elementos cromatínicos (nucleosomas). Para ello, el patrón de digestión de la cromatina con nucleasas (p. e. MNasas) permitirá comprobar

que la organización cromatínica del minicromosoma no se altera tras la inserción del elemento. Asimismo, será necesario diseñar el sitio de inserción con dianas de restricción adecuadas. Esto permitirá insertar distintas secuencias de ADN mediante los adaptadores oportunos y poder comparar sistemáticamente el ΔLk estabilizado por distintos nucleosomas o elementos de la cromatina.

A parte de calcular los valores de ΔLk de elementos insertados individualmente, esta estrategia permitiría examinar también múltiples elementos simultáneamente. Para ello es necesario que el tamaño del sitio de unión (longitud en pb) de esos elementos sea comparable, como es el caso de los nucleosomas. Dado que el genoma de *S. cerevisiae* posiciona cerca de 60.000 nucleosomas distintos, puede generarse una librería de todas las secuencias de ADN mononucleosomales (Figura I38). Una vez insertada dicha librería en minicromosomas circulares y recolectados miles de los correspondientes clones de *S. cerevisiae*, sería factible examinar la topología de todas las moléculas de ADN (*pool*) simultáneamente en una única calle de electroforesis. En este caso, en lugar de observar una sola distribución (escalera) de topoisómeros de Lk , se observaría una distribución continua (difusa) debido a la superposición de miles de distribuciones (una por cada nucleosoma de la librería). Lo mismo ocurriría con las moléculas circulares de ADN relajado *in vitro*. El hecho de que las escaleras de topoisómeros de cada clon no queden necesariamente alineadas entre ellas y generen un patrón difuso es debido a que el tamaño de los insertos de ADN mono-nucleosomal no será idéntico en cada constructo. Sin embargo, el Lk medio de cada distribución prácticamente no se ve afectado por pequeñas diferencias de longitud entre los constructos (p. e. ± 20 pb). Esto implica que calculando el Lk medio de todas las distribuciones de la librería de minicromosomas *in vivo* y el Lk medio de las correspondientes moléculas de ADN *in vitro*, se obtiene el valor de ΔLk de los minicromosomas con la librería ($\Delta Lk^{CRM+lib}$). Este valor podrá compararse entonces con el ΔLk del minicromosoma sin la librería (ΔLk^{CRM}), de la misma manera que se realiza para un solo elemento. Por lo que, $\Delta \Delta Lk^{lib} = \Delta Lk^{CRM+lib} - \Delta Lk^{CRM}$. El valor $\Delta \Delta Lk^{lib}$ nos indicaría

el valor de Lk medio que estabiliza la librería de nucleosomas de *S. cerevisiae* *in vivo* (Figura I38).

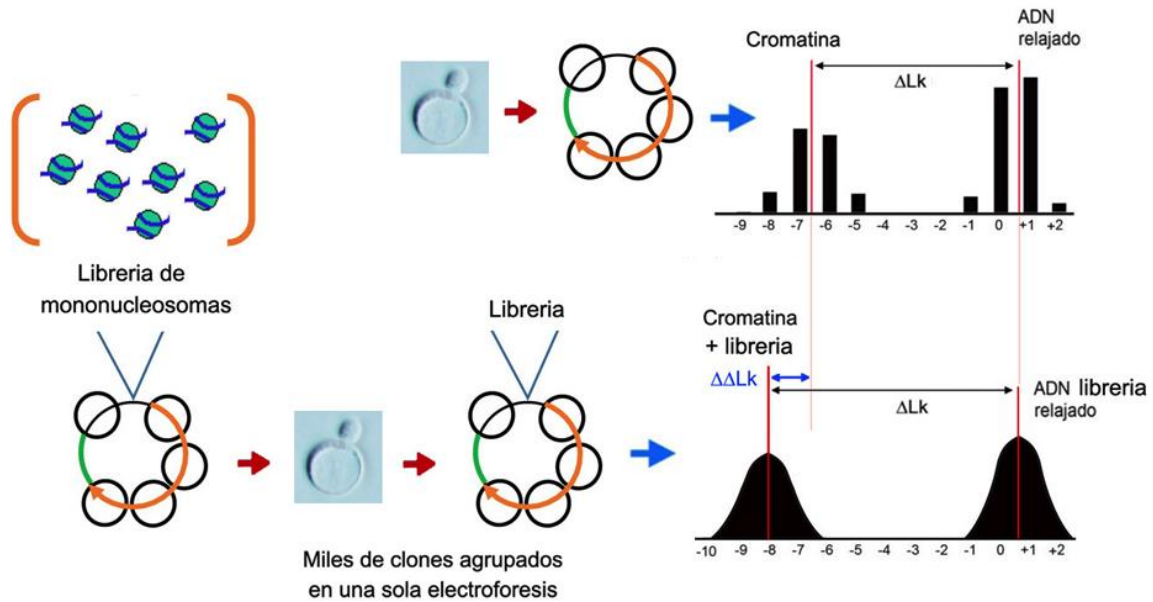


Figura I38. Cálculo del ΔLk medio estabilizado por una librería de nucleosomas. Los fragmentos mononucleosomales obtenidos mediante digestión con MNasa se insertan en minicromosomas circulares. Al agrupar todos los clones de la librería en un único carril de la electroforesis se obtendrá una distribución continua formada por la superposición de todas las distribuciones de Lk . De la misma manera que para un único elemento, el ΔLk medio del *pool* respecto a sus correspondientes ADNs relajados se compara con el ΔLk del minicromosoma sin librería para obtener el valor estabilizado por los nucleosomas de la librería: $\Delta\Delta Lk^{lib} = \Delta Lk^{CRM+lib} - \Delta Lk^{CRM}$.

objetivos

1. Diseñar minicromosomas circulares para analizar el valor ΔLk estabilizado por distintos elementos estructurales de la cromatina en *S. cerevisiae*.
2. Construir una librería de nucleosomas de *S. cerevisiae* e insertarla en minicromosomas circulares.
3. Determinar el valor medio de ΔLk estabilizado por los nucleosomas de *S. cerevisiae* *in vivo*.

publicaciones

INFORME DEL DIRECTOR DE TESIS

Artículo #1

DNA Topology and Global Architecture of Point Centromeres.

Ofelia Díaz-Ingelmo, Belén Martínez-García, Joana Segura, Antonio Valdés, and Joaquim Roca. **Cell Reports** 13, 667-677 (2015). DOI: 10.1016/j.celrep.2015.09.039

En este primer artículo se alcanza el objetivo número 1 de esta tesis. Joana Segura trabaja en el diseño y construcción de minicromosomas circulares para estudiar la topología de elementos estructurales de la cromatina. Su contribución se centra especialmente en la reducción del tamaño del minicromosoma *TRP1ARS1* y en el desarrollo de los sistemas de análisis electroforéticos para calcular valores ΔLk . Con esta estrategia se consigue descifrar la topología del centrómero puntual de *S. cerevisiae*. Los resultados sugieren un nuevo modelo estructural, que explica muchas observaciones que parecían incoherentes hasta el momento. El impacto de este trabajo lo acreditan, por ejemplo, el artículo "Preview" que se publicó en el mismo número de *Cell Reports* y los comentarios publicados por expertos de *F1000 Faculty* (F1000Prime).

A Positive Twist to the Centromeric Nucleosome

Josefina Ocampo and David J. Clark.

Preview in Cell Reports 13, October 27, 2015

"Centromeric nucleosomes are critical for chromosome attachment to the mitotic spindle. In this issue of Cell Reports, Diaz-Ingelmo et al. (2015) propose that the yeast centromeric nucleosome is stabilized by a positively supercoiled loop formed by the sequence-specific CBF3 complex".

F1000Prime - Cell Biology January 22, 2016

Yamini Dalal and Daniël Melters

GOOD FOR TEACHING | INTERESTING HYPOTHESIS | NEW FINDING

"This paper teases apart the mechanistic details of how the three elements (CDEI, II and III) at the short genetic centromere of budding yeast contribute to changes in DNA topology. By adding either an intact centromere or an inverted centromere on the minichromosome, the authors observed increased positive linking number unit difference per intact CDE added, compared to the addition of a canonical nucleosome-positioning sequence. However, in an elegant series of experiments, when they mutated either CDEI, II or III individually, or mutated a combination thereof, they discovered that only CDEIII is responsible for the +2 unit difference in linkage number, not CDE I or II. CDEIII is bound by the CBF3 complex in vivo, not by CSE4 directly. In order to explain how the CSE4 nucleosome can tolerate positive supercoils that are present on the intact centromere, the authors modeled the structure of the centromere with the CBF3 complex and found that this complex induces a positively constrained 70bp loop precisely where CSE4 nucleosome has been mapped. Altogether, this work is a beautiful mechanistic dissection of previous reports of positively supercoiled CENP-A/CSE4 nucleosomes and provides insights into the formation of discrete localized topological domains in vivo".

Artículo #2

Intracellular nucleosomes constrain a DNA linking number difference of -1.26 that reconciles the Lk paradox.

Joana Segura, Ricky S. Joshi, Ofelia Díaz-Ingelmo, Antonio Valdés, Silvia Dyson, Belén Martínez-García and Joaquim Roca.

Nature Com Sep 28;9(1):3989 (2018). DOI: 10.1038/s41467-018-06547-w

En este segundo artículo se alcanzan los objetivos 2 y 3 de esta tesis. Joana Segura completa el diseño de minicromosomas circulares pequeños, en los que incluye un centrómero y un sitio con las dianas adecuadas para la inserción de elementos estructurales de la cromatina. Su investigación se centra después en construir una librería de fragmentos de ADN mononucleosomal. Para ello, repara los extremos de fragmentos obtenidos tras la digestión de cromatina de *S. cerevisiae* y los liga a adaptadores que diseña *ad hoc*. Joana introduce la librería en un plásmido nodriza que amplifica en *E. coli*. A partir de los plásmidos de *E. coli*, escinde y circulariza los minicromosomas que contienen la librería. Introduce después los minicromosomas en *S. cerevisiae* y recolecta miles de colonias. Joana extrae minicromosomas de varias colonias individuales y agrupa el resto de colonias en una única muestra. Joana confirma la organización nucleosómica de los minicromosomas mediante digestión con nucleasa micrococcal y comprueba la representatividad de la librería de mononucleosomas mediante secuenciación masiva de los fragmentos insertados. Joana lleva a cabo, entonces, el análisis electroforético de minicromosomas individuales y calcula el ΔLk estabilizado por varios nucleosomas individuales. Finalmente, realiza el análisis electroforético de la muestra que contiene todos los minicromosomas de la librería. Consigue calcular así por primera vez el valor de ΔLk medio estabilizado por un gran número de nucleosomas en cromatina nativa.

Los resultados de este largo trabajo permiten aclarar un misterio que había sido objeto de debate durante cuatro décadas. Las revisiones destacan también su innovación

tecnológica y calidad experimental. Su relevancia se acredita, por ejemplo, en los comentarios publicados por expertos de *F1000 Faculty* (F1000Prime).

F1000Prime - Chemical Biology 29 November 2018

Tom Tullius

CONTROVERSIAL | NEW FINDING | TECHNICAL ADVANCE

"This paper revisits a longstanding conundrum in DNA structure - the so-called 'linking number paradox' - which has been discussed for more than 40 years. The essence of the paradox is that while DNA is known to wrap twice around the histone octamer in the nucleosome core particle, which would produce a theoretical change in linking number of around $\cdot 2$, experimental measurements of the change in linking number imparted by a set of nucleosomes in a minichromosome gave values near $\cdot 1$. Further experiments demonstrated that the helical twist of DNA differs when it is in the nucleosome core particle versus when it is free in solution, which explains some but not all of the paradox. Here the authors perform a series of carefully controlled experiments to measure the change in linking number imparted by nucleosome formation, and come up with a consistent value of around -1.26 (instead of the value of -1 previously assumed from early measurements). By using their new value for the change in linking number, along with well accepted values for the change in helical twist and in writhe that occur upon nucleosome formation, the authors make the linking number paradox disappear"

Firmado: Joaquim Roca Bosch

Director de la tesis

artículo #1

RESUMEN ARTÍCULO #1

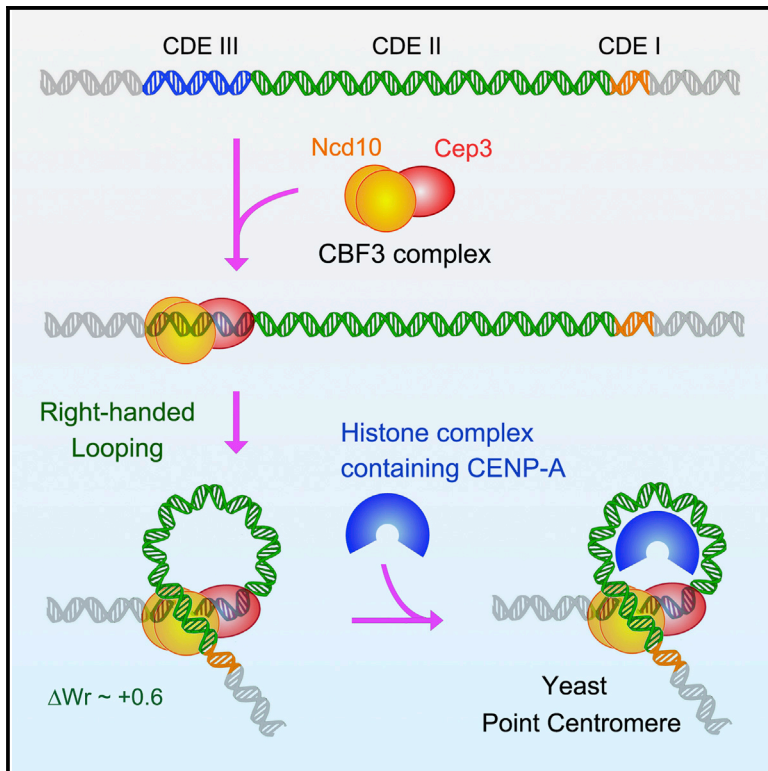
“DNA Topology and Global Architecture of Point Centromeres”

Ofelia Díaz-Ingelmo, Belén Martínez-García, **Joana Segura**, Antonio Valdés, Joaquim Roca.

Los nucleosomas canónicos se caracterizan por presentar el ADN enrollado en sentido levógiro sobre el octámero de histonas. De la misma manera, aquellos nucleosomas que incorporan la histona centromérica CENP-A, variante de la histona H3, presentan el ADN en sentido levógiro cuando son ensamblados *in vitro* sobre un octámero de histonas. Sin embargo, estudios *in vivo* sugieren que el ADN se encuentra enrollado en sentido dextrógiro sobre el nucleosoma que conforma el centrómero puntual de *Saccharomyces cerevisiae*. En este trabajo, hemos determinado el valor de la diferencia del número de enlace (ΔLk) estabilizada por el centrómero puntual y la contribución de cada uno de los elementos que lo conforman (CDEI, CDEII y CDEIII). Mediante minicromosomas circulares con y sin centrómero, hemos fijado el Lk del minicromosoma *in vivo* (Lk^{ch}) y se ha determinado, de forma exacta, el valor de ΔLk relativo al minicromosoma relajado (Lk_0). Los resultados muestran que el nucleosoma centromérico estabiliza un valor de $\Delta Lk = +0.6$ unidades. Esta topología depende de los elementos CDEII y CDEIII y, en cambio, es independiente tanto de la unión del factor Cbf1 al elemento CDEI como de la longitud variable de CDEII. Estos resultados son consistentes con un modelo del centrómero puntual donde el complejo CBF3 forma un bucle de ADN (*loop*) dextrógiro que incluye los segmentos CDEII y CDEIII. Posteriormente, este *loop* dextrógiro se estabilizaría por un hemisoma de histonas que incluye CENP-A.

DNA Topology and Global Architecture of Point Centromeres

Graphical Abstract



Authors

Ofelia Díaz-Ingelmo, Belén Martínez-García, Joana Segura, Antonio Valdés, Joaquim Roca

Correspondence

joaquim.roca@ibmb.csic.es

In Brief

Díaz-Ingelmo et al. show that the point centromere of budding yeast constrains a positive DNA supercoil, which requires the centromere DNA elements CDEII and CDEIII, but not CDEI. This positive supercoil suffices to accommodate a sub-octameric histone complex, which does not have to be inherently right-handed.

Highlights

- Point centromeres increase the ΔLk of DNA by +0.6 units in yeast minichromosomes
- This ΔLk change is constrained by the topology of CDEs within the point centromere
- Centromere DNA topology is independent of CDEII length and of *cbf1* binding to CDEI
- Constraining of ΔLk +0.6 is established by the integrity of CDEII and CDEIII



DNA Topology and Global Architecture of Point Centromeres

Ofelia Díaz-Ingelmo,¹ Belén Martínez-García,¹ Joana Segura,¹ Antonio Valdés,¹ and Joaquim Roca^{1,*}¹Molecular Biology Institute of Barcelona (IBMB), Spanish National Research Council (CSIC), Barcelona 08028, Spain*Correspondence: joaquim.roca@ibmb.csic.es<http://dx.doi.org/10.1016/j.celrep.2015.09.039>This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

SUMMARY

DNA is wrapped in a left-handed fashion around histone octasomes containing the centromeric histone H3 variant CENP-A. However, DNA topology studies have suggested that DNA is wrapped in a right-handed manner around the CENP-A nucleosome that occupies the yeast point centromere. Here, we determine the DNA linking number difference (ΔLk) stabilized by the yeast centromere and the contribution of the centromere determining elements (CDEI, CDEII, and CDEIII). We show that the intrinsic architecture of the yeast centromere stabilizes +0.6 units of ΔLk . This topology depends on the integrity of CDEI and CDEIII, but it is independent of cbf1 binding to CDEI and of the variable length of CDEII. These findings suggest that the interaction of the CBF3 complex with CDEIII and a distal CDEII segment configures a right-handed DNA loop that excludes CDEI. This loop is then occupied by a CENP-A histone complex, which does not have to be inherently right-handed.

INTRODUCTION

The centromere is the genetic locus that organizes the kinetochore, the multi-protein complex that attaches each chromosome to spindle microtubules during mitosis and meiosis (Biggins, 2013; Bloom, 2014). In contrast to most eukaryotic centromeres that are epigenetically defined and span large domains of heterochromatin (Black et al., 2010; Henikoff and Furuyama, 2010; Burrack and Berman, 2012), the centromere of budding yeast is a compact structure occupying ~200 bp of DNA (Bloom and Carbon, 1982) and is genetically defined by three conserved centromere-determining elements (CDEs): an 8-bp palindrome called CDEI, a 26-bp sequence called CDEIII, and a 78- to 86-bp stretch of AT-rich (~90%) DNA called CDEII that lies in between (Clarke, 1998). Despite these differences in size and locus definition, all eukaryotic centromeres have a common protein determinant, namely, the histone H3 variant CENP-A (also known as Cid in *Drosophila* and Cse4 in budding yeast). Most eukaryotic centromeres thus comprise hundreds of nucleosomes that contain CENP-A (Schueler and Sullivan, 2006), whereas budding yeast contains fewer CENP-A nucleosomes (Haase

et al., 2013), including a single CENP-A nucleosome positioned on the CDEs (Furuyama and Biggins, 2007; Cole et al., 2011; Krassovsky et al., 2012).

Although CENP-A nucleosomes are essential for kinetochore formation, their functional properties are unknown. Numerous lines of evidence indicate that their structure differs from canonical nucleosomes. Regarding their histone composition and CENP-A copy number, several mutually exclusive structures have been proposed. These models include conventional (CENP-A /H4/H2B/H2A)₂ octasomes (Camahort et al., 2009; Zhang et al., 2012; Padeganeh et al., 2013; Wisniewski et al., 2014), asymmetric (CENP-A/H3/(H4/H2B/H2A)₂ octasomes (Lochmann and Ivanov, 2012), (CENP-A/H4/Scm3)₂ hexasomes (Mizuguchi et al., 2007), (CENP-A/H4)₂ tetrasomes (Xiao et al., 2011; Aravamudhan et al., 2013), and CENP-A/H4/H2B/H2A hemisomes (Henikoff and Furuyama, 2012; Furuyama et al., 2013). Regarding their morphology, atomic force microscopy (AFM) studies showed that in-vitro-assembled CENP-A nucleosomes have a reduced height (Dalal et al., 2007; Dimitriadis et al., 2010; Bui et al., 2012), which is established by the CATD domain of CENP-A (Miell et al., 2013). At the point centromere of budding yeast, a histone complex containing CENP-A interacts with CDEII and is flanked by the proteins that bind to CDEI and CDEIII (Krassovsky et al., 2012). CDEI is occupied by the general transcription factor Cbf1(p39), which is not essential for centromere function (Cai and Davis, 1989; Baker et al., 1989). CDEIII is occupied by the CBF3 complex, which contains four essential proteins, namely, a Cep3 (p64) homodimer, an Skp1(p19)-Ctf13(p58) heterodimer, and an Ndc10 (p110) homodimer (Jiang et al., 1993; Lechner and Carbon, 1991; Connelly and Hieter, 1996). The observation of physical interactions between CBF3 and Cbf1 (Hemmerich et al., 2000) has led to the proposal that CDEI and CDEIII are bridged to hold a CDEII loop that stabilizes the CENP-A nucleosome (Xiao et al., 2011; Cho and Harrison, 2011).

A second striking feature of centromeric nucleosomes regards their DNA topology. Early studies described that, in yeast, circular minichromosomes with and without a centromere have unequal distributions of DNA topoisomers (Bloom et al., 1983, 1984). More recent analyses revealed that in vitro chromatin assembly with CENP-A/H4/H2B/H2A and the histone chaperone RbAp48 from *Drosophila* tends to constrain positive DNA supercoils, in contrast to the negative supercoils stabilized by conventional nucleosomes (Furuyama and Henikoff, 2009). The same study showed that the presence of a point centromere instead of a regular nucleosome in yeast circular minichromosomes

leads to a difference of approximately +2 units in the linking number of DNA (Lk). A similar Lk deviation was observed with the partitioning locus of the yeast 2- μ m plasmid, which also includes a CENP-A nucleosome (Huang et al., 2011). On the basis of these observations, it was postulated that DNA wraps in a right-handed orientation in CENP-A nucleosomes (Furuyama and Henikoff, 2009; Henikoff and Furuyama, 2012). However, crystal and biochemical studies of CENP-A nucleosomes reconstituted in vitro demonstrate that they are left-handed and restrain negative DNA supercoils in a similar way to canonical octasomes (Sekulic et al., 2010; Tachiwana et al., 2011). This inconsistency has been exploited to support the notion that CENP-A nucleosomes are not left-handed octasomes (Henikoff and Furuyama, 2012). However, there is no experimental evidence of inherently right-handed histone complexes. The Lk differences observed in yeast centromeric minichromosomes could be explained by other protein-DNA interactions (Mishra et al., 2013) or caused by alterations of DNA topology outside the point centromere, induced either in vivo or during experimental manipulation.

Here, we analyzed the topology of DNA at the point centromere of budding yeast. Unlike previous studies that directly compared the Lk of circular minichromosomes with and without a functional centromere, we fixed the Lk of the minichromosomes in vivo (Lk^{ch}) and determined their exact Lk difference relative to relaxed DNA circles (Lk^o). Next, we minimized the chromatin structure of the minichromosomes to discern whether the Lk deviations were constrained by the point centromeres or were instead due to alterations of the adjacent chromatin. Finally, we examined the contribution of CDEs in determining centromere DNA topology. Our results revealed that the intrinsic architecture of the point centromere stabilizes an Lk difference of +0.6 and that this topology is configured by the protein complexes bound to CDEIII and CDEII, but not to CDEI. These findings support a model of the point centromere in which the CBF3 complex configures a right-handed loop of DNA that includes the CDEIII and CDEII segments. This loop is then occupied by a CENP-A histone complex, which does not have to be innately right-handed.

RESULTS

The DNA Linking Number Difference in the Yeast TA1 Minichromosome

Circularization of the EcoRI 1,453-bp genomic fragment of *Saccharomyces cerevisiae*, which comprises the TRP1 gene and the ARS1 origin, generates a minichromosome (TA1), which replicates and segregates as a multicopy episome in *trp1* yeast strains (Thoma et al., 1984). TA1 maintains the nucleosome organization of the genomic loci (Thoma et al., 1984; Jiang and Pugh, 2009). Four nucleosomes (I–IV) are positioned downstream of the transcription start site of TRP1 and three nucleosomes (V–VII) downstream of the ARS1 region. We performed micrococcal nuclease digestions to confirm that the positions of these seven nucleosome are preserved in the TA1 minichromosome of our yeast cells (Figures 1A and S1).

As eukaryotic nucleosomes constrain negative DNA supercoils, the DNA linking number of circular minichromosomes (Lk^{ch}) is reduced with respect to that of relaxed DNA (Lk^o)

by a value that roughly correlates with the number of nucleosomes assembled (Prunell, 1998). For the purpose of our study, we sought to accurately determine the Lk difference ($\Delta Lk = Lk^o - Lk^{ch}$) of TA1. Since Lk^o changes with temperature (Depew and Wang, 1975), we relaxed naked DNA circles at the same temperature that the yeast cultures. Likewise, since Lk^{ch} depends also on temperature and it can be altered by endogenous topoisomerases during cell disruption and DNA extraction, we fixed the in vivo Lk^{ch} values by quenching the yeast cultures with a cold (-20°C) ethanol-toluene solution. This fixation step irreversibly inactivates the cellular topoisomerases, so precluding the alteration of the in vivo Lk^{ch} values in subsequent manipulations (Figure S2).

Following the above considerations, we examined by gel electrophoresis the Lk distributions of TA1 fixed in vivo at 26°C and that of the naked TA1 circle relaxed in vitro at 26°C . In order to resolve the topoisomers of both Lk distributions in a single one-dimensional gel, we adjusted the concentration of chloroquine to 0.2 $\mu\text{g/ml}$ during electrophoresis (Figure 1B). Otherwise, in absence of chloroquine, the Lk distribution of the negatively supercoiled DNA extracted from the minichromosomes would be collapsed in a fast-migration band and the Lk distribution of the relaxed DNA would overlap with the nicked circles (Figure S3). We examined also the same DNA samples in a two-dimensional gel along with a marker of Lk topoisomers, in order to unambiguously count the number of Lk topoisomers that separate both Lk distributions (Figure 1C). After counting the topoisomer bands and quantifying their individual intensities, we calculated ΔLk as the distance (Lk units) between the midpoints of the two distributions (Lk^o and Lk^{ch}). Measurements in four independent yeast cultures indicated that TA1 has an in vivo $\Delta Lk = -9.4$ (Figure 1D).

Next, we sought to discern whether the ΔLk of TA1 was constrained by its chromatin structure. For this purpose, we cultured the same yeast cells skipping the fixation step and lysed them in order to solubilize TA1. We added a negatively supercoiled plasmid (internal DNA control) and catalytic amounts of topoisomerases (vaccinia virus topoisomerase [topo] I and *S. cerevisiae* topoisomerase II) to the lysate. We incubated the mixtures at 26°C to allow relaxation of free DNA supercoils. Electrophoretic analyses of these samples revealed that the Lk distribution of the solubilized TA1 minichromosomes was nearly identical to that of samples fixed in vivo and that it was not appreciably altered after incubation with topoisomerase I or topoisomerase II (Figure 1E, top). Conversely, the control plasmid included in the reactions became near fully relaxed by the endogenous topoisomerases present in the yeast lysate and completely relaxed following the addition of topoisomerase I or topoisomerase II (Figure 1E, bottom). This result indicated that the -9.4 units of ΔLk in the TA1 minichromosome are stably constrained by its chromatin structure.

A Point Centromere Deviates the Lk Difference of Circular Minichromosomes by +0.6 Units

Having established that TA1 stabilizes a ΔLk of -9.4 , we examined how this ΔLk was modified after inserting additional chromatin elements into the minichromosome. All these experiments, including the in vivo fixation of Lk values, the incubation of

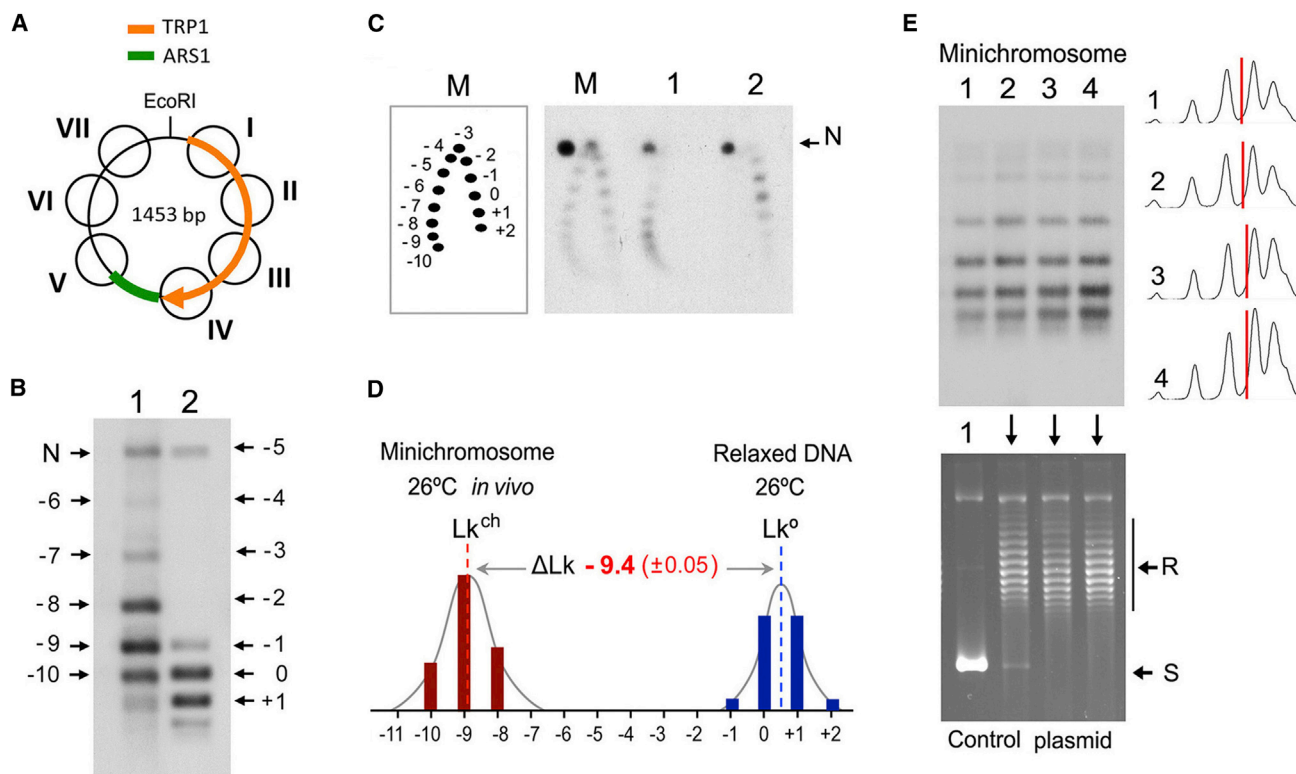


Figure 1. DNA Linking Number Difference in the Yeast TA1 Minichromosome

(A) Scheme of the TA1 minichromosome indicating the EcoRI circularization site and the position of nucleosomes I to VII with respect to the TRP1 gene and ARS1 region as determined by Thoma et al. (1984) and confirmed in the present study (see also Figure S1).

(B) One-dimensional gel electrophoresis of the distribution of Lk topoisomers of the TA1 minichromosome fixed in vivo at 26°C (lane 1) and of the naked TA1 circle relaxed in vitro with topoisomerase I at 26°C (lane 2). Electrophoresis was in the presence of 0.2 μg/ml chloroquine as explained in Figure S3.

(C) Two-dimensional gel electrophoresis of the same DNA samples (lanes 1 and 2) with a marker of Lk topoisomers (lane M). In both one- and two-dimensional gels (B and C), individual Lk topoisomers are identified by correlative numbers (not Lk values) starting with 0 at the main topoisomer of the relaxed DNA and decreasing toward the minichromosome DNA. N, nicked circles.

(D) Histogram of the signal intensity of individual Lk topoisomers displayed in the above gel blots. The ΔLk of the minichromosome is the distance (Lk units) between the midpoint of the in vivo Lk distribution (Lk^{ch}) and the midpoint of the Lk distribution of the relaxed DNA (Lk^o). The mean ΔLk (±SD) of four independent experiments is indicated.

(E) The gel on the top compares the Lk distributions of the TA1 minichromosome extracted from fixed cells at 26°C (lane 1), solubilized in cell lysates at 26°C (lane 2), and after incubation with topoisomerase I (lane 3) or topoisomerase II (lane 4). Plots of the gel lanes and the center of intensities are shown. The gel on the bottom shows the topology of the negatively supercoiled plasmid (lane 1) that was added in excess to the above incubations (lanes 2, 3, and 4) as internal control for the relaxation activity of topoisomerases. S, supercoiled forms. R, relaxed Lk distribution.

solubilized minichromosomes, and the relaxation of naked DNA, were done at the same temperature (26°C). We compared the effect of three inserts, which contained the following: the positioning sequence of the nucleosome *High2* of *S. cerevisiae* (Segal et al., 2006), the point centromere of yeast chromosome IV (*CEN4*), and *CEN4* with a mutated CDEIII sequence that precludes centromere assembly (*mutCEN*) (Jehn et al., 1991). The three inserts were of identical length and were located upstream of the TRP1 gene (between nucleosomes I and VII). The result was three circular minichromosomes of 1686 bp (Figures 2A and S4).

As expected, upon yeast transformation, minichromosomes with the functional *CEN4* centromere were recovered in low copy number with respect to those with the *High2* and *mutCEN* inserts, which were recovered as high-copy episomes. As for TA1, we fixed the topology of these minichromosomes in vivo,

examined their Lk distributions in one- and two-dimensional gels (Figure 2B), and calculated their ΔLk with respect to relaxed DNA circles of 1,686 bp (Figure 2C). We found that the minichromosome with the *High2* insert had a ΔLk of −10.8. This value implied a gain of −1.4 Lk units relative to TA1 and was thus compatible with the assembly of *High2* and other protein-DNA complexes at the inserted segment. Likewise, the minichromosome containing *mutCEN* had a ΔLk of −10.6, a gain of −1.2 Lk units relative to TA1, which also fitted the assembly of at least one additional nucleosome. Conversely, the minichromosome with *CEN4* presented a ΔLk of −8.8. This value implied a gain of +0.6 Lk units relative to TA1 and of +2 Lk units with respect to the minichromosome with the *High2* insert. This Lk change was incompatible with the accommodation of a left-handed nucleosomal particle at the *CEN4* sequence, unless other regions of the minichromosome had markedly altered their DNA

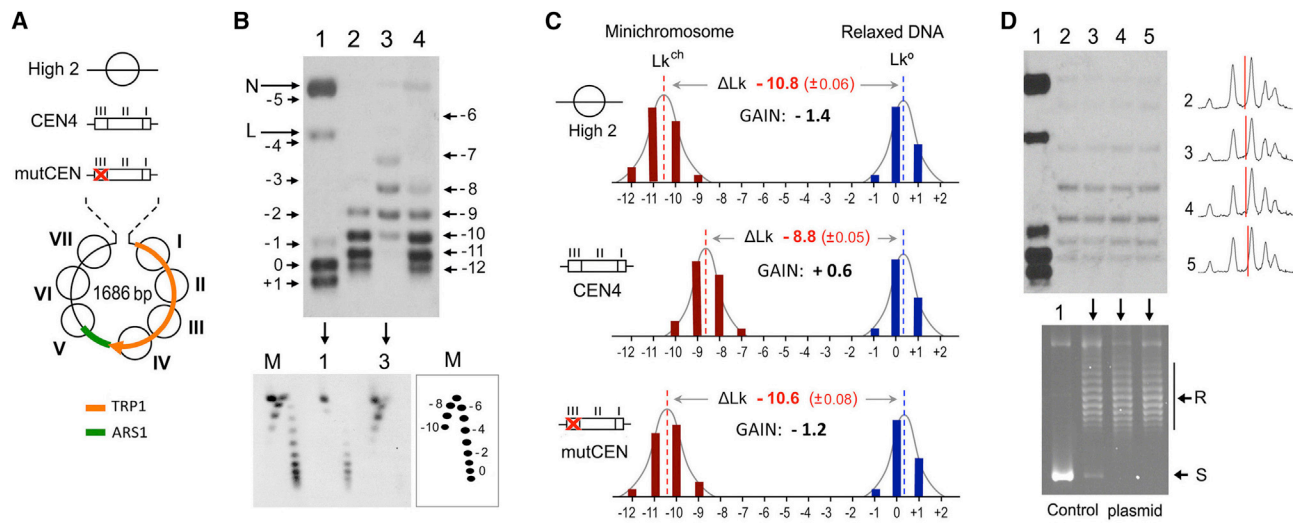


Figure 2. Lk Differences of Centromeric and Non-centromeric Minichromosomes of 1,686 bp

(A) Scheme of the minichromosomes of 1686 bp that were generated following the insertion of the *High2*, *CEN4*, and *mutCEN* sequences in-between the nucleosomes I and VII of TA1. See Figure S4 for details. The orientation of CDEs of *CEN4* is indicated (I, II, and III).

(B) The gel-blot on top shows the *Lk* distributions of 1,686-bp DNA circles relaxed at 26°C (lane 1), and of 1686 bp minichromosomes fixed in vivo at 26°C and that contained the *High2* (lane 2), *CEN4* (lane 3), and *mutCEN* (lane 4) inserts. One-dimensional electrophoresis was done as in Figure 1B. N, nicked circles. L, linearized circles. The two-dimensional gel (bottom) shows the *Lk* distributions of the same samples of relaxed DNA (lane 1) and of the *CEN4* minichromosome (lane 3) with a marker of *Lk* topoisomers (lane M). In both gel blots, individual *Lk* topoisomers are identified by correlative numbers.

(C) Aligned histograms of the *Lk* distributions of minichromosomes with the *High2*, *CEN4*, and *mutCEN* inserts and of the relaxed DNA circles. ΔLk values were calculated as in Figure 1. The mean ΔLk (\pm SD) from three independent experiments is indicated. The gain value produced by each insert is the $\Delta \Delta Lk$ with respect to the ΔLk of the TA1 minichromosome ($\Delta Lk -9.4$).

(D) *Lk* distributions of the relaxed *CEN4* DNA circle (lane 1) and of the *CEN4* minichromosome fixed in vivo at 26°C (lane 2), solubilized at 26°C (lane 3), and incubated with topoisomerase I (lane 4) and topoisomerase II (lane 5). Plots of lanes 2–5 are shown. The gel on the bottom shows the control plasmid (lane 1) after its incubation with the minichromosomes of lanes 3, 4, and 5. S, supercoiled forms. R, relaxed *Lk* distribution.

topology. The ΔLk of the *CEN4* minichromosome remained unchanged when the chromatin was solubilized and incubated with topoisomerase I or topoisomerase II (Figure 2D). Therefore, the gain of +0.6 *Lk* units induced by *CEN4* was stably constrained by the minichromosome structure.

The Gain of +0.6 *Lk* Units Is an Intrinsic Trait of the Point Centromere Architecture

The gain of +0.6 *Lk* units observed in the *CEN4* minichromosomes may be stabilized by the architecture of the point centromere, but may also result from alterations of DNA topology in neighboring regions. To distinguish these two scenarios, we constructed a new set of TA1-derived minichromosomes in which we minimized the chromatin structure and the distance between functional elements (Figures 3A and S5). First, we removed nucleosomes VI and VII of TA1. We maintain nucleosome V as it partially overlaps with the ARS1 region, and we kept a minimal promoter upstream of the transcription start site of TRP1 at the 5' flank of nucleosome I. Next, we inserted the *High2*, *CEN4*, and *mutCEN* sequences as narrowly as possible in-between nucleosomes I and V, but without interfering with the TRP1 and ARS1 functions. The result was a set of stable circular minichromosomes of 1,286 bp, in which the positioning sequence of *High2* was 39 bp from the edge of nucleosome V and 67 bp from the transcription start site of TRP1. In the case of the centromere, we allocated the CDEs of *CEN4* (111 bp) in

both orientations. The centromere was functional when CDEI was 48 bp from the edge of nucleosome V and CDEIII was 93 bp from the TRP1 start site. In the reverse orientation, the centromere was functional when CDEIII was 68 bp from the edge of nucleosome V and CDEI was 73 bp from the TRP1 start site. We performed micrococcal nuclease digestions, which confirmed that the assembly of the centromere in the 1,286-bp minichromosomes did not induce loss of the flanking nucleosomes (Figure 3B). We did not obtain transformants with constructs including shorter linker regions, possibly because centromere assembly interfered with ARS1 or TRP1 functions.

We examined then the in vivo *Lk* distributions of the above 1,286-bp minichromosomes fixed in vivo (Figure 3C) and after their incubation with topoisomerases (Figure 3D). The topology of the minichromosomes was not affected by topoisomerases, which denoted that their in vivo ΔLk was stabilized by chromatin structure. With respect to relaxed DNA of the same length, the minichromosomes with the *High2* and *mutCEN* inserts presented a ΔLk of -8.2 and -8.0 , respectively, whereas the minichromosomes with *CEN4* inserted in either orientation presented a ΔLk of -6.2 (Figure 3E). Interestingly, all these ΔLk values (-8.2 , -8.0 , and -6.2) differed by 2.6 units from the ΔLk values of the minichromosomes of 1,686 bp harboring the corresponding *High2*, *mutCEN*, and *CEN4* inserts (-10.8 , -10.6 , and -8.8). This decrease in ΔLk was consistent with the loss of the negative DNA supercoils that were constrained by nucleosomes VI and VII

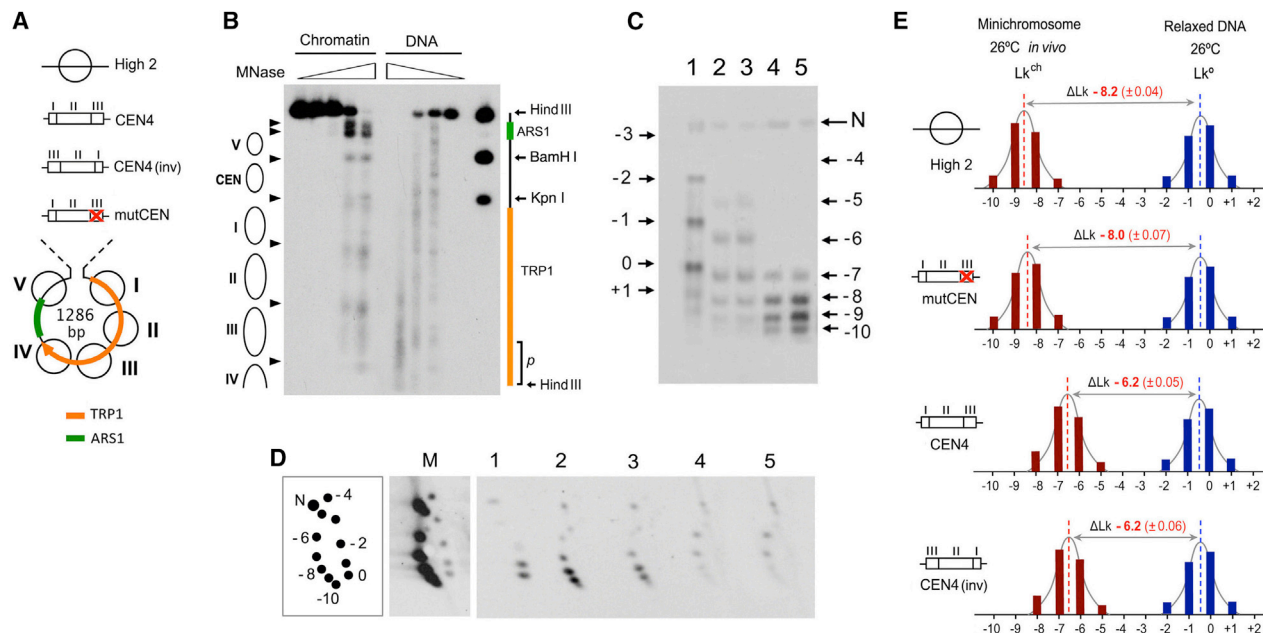


Figure 3. *Lk* Differences of Centromeric and Non-centromeric Minichromosomes of 1,286 bp

(A) Scheme of the minichromosomes of 1286 bp that were generated following the insertion of the *High2*, *CEN4*, *CEN4(inv)*, and *mutCEN* sequences closely in between nucleosomes I and V of TA1. See Figure S5 for details.

(B) Micrococcal nuclease digestion patterns of 1,286-bp minichromosomes carrying *CEN4* (chromatin) compared to naked DNA molecules (DNA). Samples digested with increasing nuclease were cut with endonuclease *HindIII*, separated on a 1.2% agarose gel, blotted, and probed with a 130-bp sequence (*p*) at the edge of the *HindIII* site.

(C) One-dimensional gel-blot comparing the *Lk* distributions of 1,286-bp DNA circles relaxed at 26°C (lane 1), and of 1286-bp minichromosomes fixed in vivo at 26°C and that contained *CEN4* (lane 2), *CEN4(inv)* (lane 3), *mutCEN* (lane 4), and *High2* (lane 5). Electrophoresis was done as in Figure 1B. N, nicked circles.

(D) Two-dimensional gel blot comparing a marker of *Lk* topoisomers (lane M) and the *Lk* distributions of 1,286-bp DNA circles relaxed at 26°C (lane 1) and of 1,286 bp minichromosomes containing *High2* (lane 2), *mutCEN* (lane 3), *CEN4* (lane 4), and *CEN4(inv)* (lane 5) following their solubilization in the presence of topoisomerase activity. N, nicked circles. In both gel blots (C and D), individual *Lk* topoisomers are identified by correlative numbers.

(E) Aligned histograms of the *Lk* distributions of the 1,286-bp minichromosomes and relaxed DNA circles. ΔLk values were calculated as in Figure 1D. Average values (\pm SD) of three independent experiments are indicated.

or nearby complexes. However, this common change of 2.6 *Lk* units also suggested that the topology of DNA in the remaining nucleosomes (I to V) was very similar in the 1,686-bp and 1,286-bp minichromosomes regardless of the elements inserted. Therefore, the gain of +0.6 *Lk* units observed in centromeric minichromosomes most likely reflects a trait of the CEN DNA topology rather than an alteration of neighboring chromatin.

The Topology of DNA at Point Centromeres Is Independent of the Length of CDEII and Is Not Altered by the Disruption of CDEI

Current models of the yeast point centromere postulate that DNA wraps around a histone core such that CDEI and CDEIII come in close proximity, and that bridging interactions between *cbf1* and CBF3 then stabilize the global complex (Hemmerich et al., 2000; Xiao et al., 2011; Cho and Harrison, 2011). However, the rotational phasing between CDEI and CDEIII is not conserved in the 16 yeast centromeres due to the variable length of CDEII (from 77 to 86 bp). Then, if CDEI and CDEIII were bridged by protein interactions, the intervening CDEII segment would have to untwist or over-twist up to about half a turn (± 0.5 *Lk* units) in some cases (Figure 4A). Consequently, the ΔLk values stabilized

by point centromeres would depend on the length of CDEII. We tested this hypothesis by comparing the topology of 1,286-bp minichromosomes that carried *CEN4*, *CEN2*, *CEN7*, and *CEN12*, which have CDEII sequences of 77, 83, 85, and 86 bp, respectively (Figures 4B and S5). We found that the four centromeres produced identical *Lk* distributions (Figure 4B), in sharp contrast to the distinct helical compensations (ΔLk relative to *CEN4*) that would be required to bridge their CDEI and CDEIII elements (Figure 4C). Therefore, the stabilization of +0.6 *Lk* units by the point centromere occurs irrespective of the length of CDEII.

Since CDEI is not essential for centromere function, we questioned then whether disruption of CDEI would affect the topology of the point centromere. For this purpose, we mutated 6 bp of the CDEI sequence of *CEN4* in the 1,286-bp minichromosome to abolish its recognition by *cbf1* (Wilmen et al., 1994) (Figure 5A and S5). The $\Delta CDEI$ minichromosome presented mitotic stability and low copy number, which corroborated that CDEI disruption does not preclude centromere function. The $\Delta CDEI$ minichromosome presented a $\Delta Lk = -6.2$, the same that of the *CEN4* minichromosome (Figures 5B and 5C). Therefore, the stabilization of +0.6 *Lk* units by point

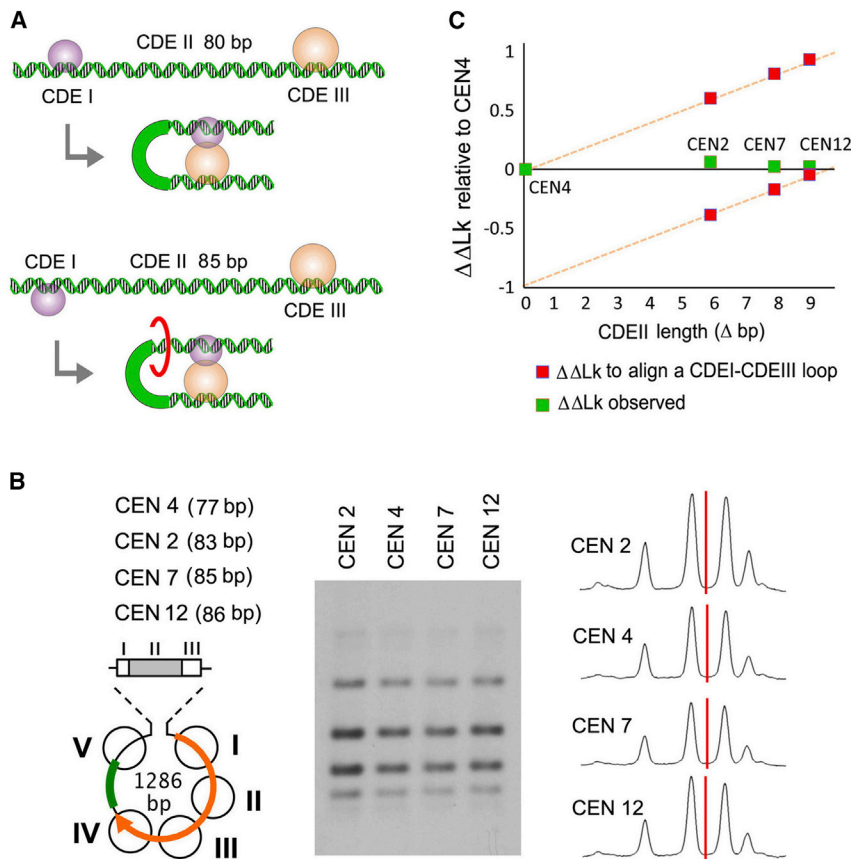


Figure 4. Effect of CDEII Length on the Topology of DNA at Point Centromeres

(A) The variable length of CDEII (e.g., 80 versus 85 bp) changes the rotational phases between CDEI and CDEIII. Bridging the proteins bound to CDEI and CDEIII would require an adjustment of the helical repeat of the intervening loop of DNA.

(B) Structure and *Lk* distributions of 1,286-bp minichromosomes that carried CEN2, CEN4, CEN7, and CEN12. The length of their corresponding CDEII element is indicated (bp). See Figure S5 for details. Electrophoresis was done as in Figure 3C. Plots of the gel lanes and the intensity center of *Lk* distribution are shown.

(C) Helical compensations (by untwisting or overtwisting) required to align CDEI and CDEIII elements as a function of CDEII length in CEN2, CEN4, CEN7, and CEN12. Theoretical values (red) are expressed as $\Delta\Delta Lk$ relative to CEN4 and compared with the experimental data (green).

centromeres was not established by protein interactions bridging CDEI and CDEIII.

We examined finally whether CDEII, which is essential for centromere function along with CDEIII, was required to determine the topology of DNA at the point centromere. For this purpose, we replaced the AT-rich CDEII sequence of *CEN4* in the 1,286-bp minichromosomes by a corresponding segment of *High2*, which is not AT rich (55% GC). We did the same replacement in the $\Delta CDEI$ construct, such that only the CDEIII element was conserved in the insert (Figures 5A and S5). The resulting $\Delta CDEII$ and $\Delta CDEI+II$ minichromosomes presented low mitotic stability and were recovered as high-copy episomes, which corroborated the loss of centromere function. Both $\Delta CDEII$ and $\Delta CDEI+II$ minichromosomes presented similar DNA topology as they had a ΔLk of -6.9 and -7.0 , respectively (Figures 5B and 5C). These ΔLk values indicated that the inserts with $\Delta CDEII$ and $\Delta CDEI+II$ do not produce the gain of $+0.6 Lk$ units observed with the full *CEN4* sequence. Therefore, both CDEIII and CDEII are necessary to establish the topology of DNA at the point centromere.

DISCUSSION

Here, we have shown that the point centromere of budding yeast stabilizes a ΔLk of $+0.6$ units and that this topology is determined by the protein complexes bound to CDEII and/or CDEIII, but not

to CDEI. To reach these conclusions, we excluded other plausible mechanisms that could account for the *Lk* differences between centromeric and non-centromeric constructs. First, we eliminated plausible alterations of *Lk* values during cell disruption and DNA extraction. We found that changes in temperature during sample manipulation markedly deviate the *Lk* of yeast minichromosomes. Thus, we fixed the *in vivo* ΔLk values by quench-

ing the yeast cultures with a procedure that irreversibly inactivates the cellular topoisomerases. Second, we solubilized the minichromosomes and incubated them with topoisomerases at the same temperature as that used for the cultures in order to discern whether *in vivo* ΔLk values were constrained by chromatin structure or were relaxable as free DNA supercoils. Third, in addition to comparing the *Lk* distributions of distinct minichromosomes directly as in previous studies (Bloom et al., 1983, 1984; Furuyama and Henikoff, 2009), we determined the ΔLk of each minichromosome using Lk° as a reference value and then compared the ΔLk of different minichromosomes. To calculate ΔLk accurately, we obtained Lk° by relaxing DNA circles of the same length and at the same temperature as used for the *in vivo* minichromosomes. Only using this ΔLk approach could we determine unambiguously the individual contribution (*Lk* gain) of the elements that we inserted or deleted from the minichromosomes. In this regard, our experimental data show that the TA1 minichromosome constraints $\Delta Lk = -9.4$. This measurement corrects the ΔLk approximately -7 of TA1 reported in earlier studies, in which ΔLk was determined relative to the gel-position nicked circles instead from Lk° (Pederson et al., 1986). Since the TA1 minichromosome accommodates seven nucleosomes and the average value of ΔLk stabilized by nucleosomes is -1 (Prunell, 1998), it is possible that other protein-DNA complexes or high-order folding of chromatin may contribute additional changes to the *Lk* of the minichromosome *in vivo*. In addition,

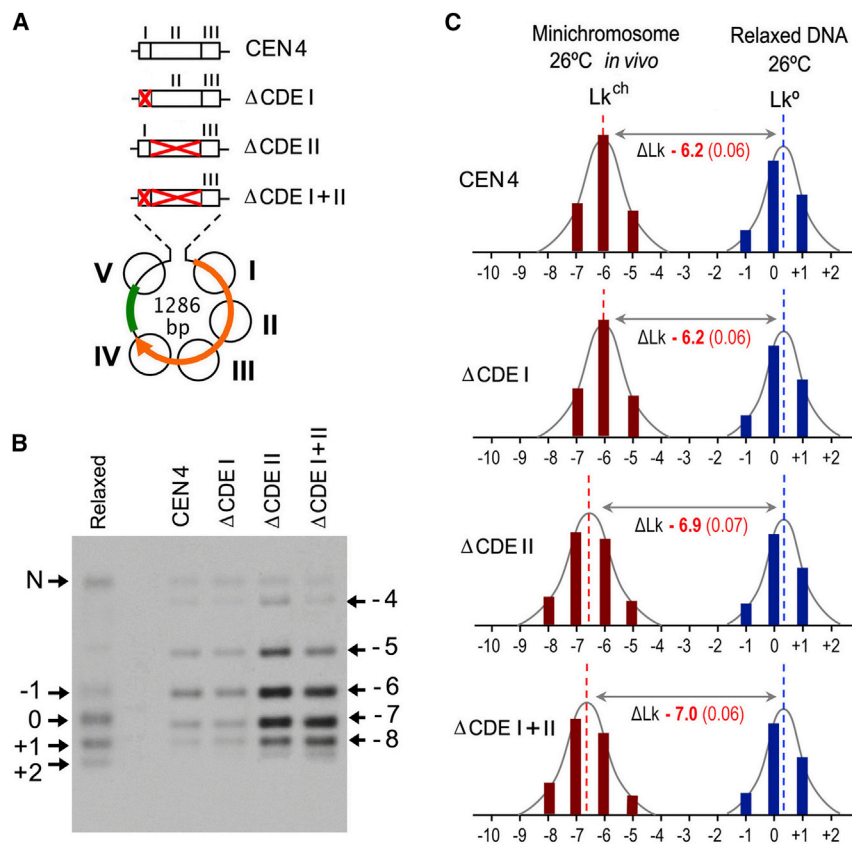


Figure 5. Effect of the Disruption of CDEI and CDEII on the Topology of DNA at Point Centromeres

(A) Structure and *Lk* distributions of 1,286-bp minichromosomes that carried *CEN4* and *CEN4* with CDEI disrupted (Δ CDEI), CDEII disrupted (Δ CDEII), and both CDEI and CDEII disrupted (Δ CDEI+II). See Figure S5 for details.

(B) *Lk* distributions of the above 1,286-bp minichromosomes and of relaxed DNA of the same length. Electrophoresis and counting of individual topoisomers was done as in Figure 3C.

(C) Aligned histograms of the *Lk* distributions of relaxed DNA circles and of the 1,286-bp minichromosomes that carried *CEN4*, Δ CDEI, Δ CDEII, and Δ CDEI+II. Δ *Lk* values were calculated as in Figure 1D. Average values (\pm SD) of three independent experiments are indicated.

given that nucleosomes are more dynamic and structurally heterogeneous than initially assumed (Zlatanova et al., 2009; Rhee et al., 2014), it is also plausible that not all nucleosomes or nucleosome-like particles produce the same Δ *Lk* values in vivo. Some complexes may induce Δ *Lk* < -1, while others may do the opposite (as it is the case of point centromeres). Following our approach, we show that the insertion of a point centromere produces always a gain of +0.6 *Lk* units, whereas an insert containing the sequence of a stable nucleosome (*High2*) and the *mutCEN* sequence produce gains of -1.4 and -1.2 *Lk* units, respectively. Our results are thus consistent with the difference of about +2 *Lk* units between centromeric and non-centromeric minichromosomes observed in previous studies (Bloom et al., 1983, 1984; Furuyama and Henikoff, 2009). Finally, we discarded that the gain of +0.6 Δ *Lk* units produced by a point centromere is due to alterations of DNA topology induced in the neighboring regions. We show that by reducing the size of the minichromosomes and the inserted sequences, there is no dysfunction of the adjacent TRP and ARS elements as long as the CDEs are flanked by minimum linker segments and that this proximity does not induce loss of nucleosomes adjacent to point centromeres. These observations are consistent with previous studies showing that the native yeast point centromeres are generally flanked by regularly spaced nucleosomes (Bloom and Carbon, 1982; Hsu et al., 2003; Furuyama and Biggins, 2007; Gkikopoulos et al., 2011; Cole et al., 2011). Therefore, the stabilization of +0.6 Δ *Lk* units is an intrinsic trait of the point centromere architecture.

Next, we have shown that the stabilization of +0.6 *Lk* units by the point centromere is independent of the variable length of CDEII and thus the rotational phasing between CDEI and CDEIII. This finding argues against the existence of tight interactions between *cbf1* and CBF3 that determine the topology of the DNA (Hemerich et al., 2000; Xiao et al., 2011; Cho and Harrison, 2011). We corroborated this conclusion by showing that CDEI mutations that abolish Cbf1 binding (Wilmen

et al., 1994) also have no effect on the DNA topology. This observation was surprising, because Cbf1 induces a strong bend ($\sim 70^\circ$) in CDEI (Niedenthal et al., 1993). Therefore, given that CDEI is not essential for centromere activity, we conclude that the CDEI-*cbf1* complex and its induced bend are irrelevant in defining the essential topology of the point centromere. In contrast, we have shown that the stabilization of the +0.6 *Lk* units depends on both the CDEII and CDEIII segments. Disruption of CDEII abolishes the stabilization of the +0.6 *Lk* units, even though CBF3 may still bind to CDEIII. Disruption of CDEIII (*mutCEN*) produces a Δ *Lk* of -1.2 *Lk* units, which suggests the assembly of a conventional nucleosome in the place of the point centromere.

The protein complexes that interact with CDEII and CDEIII may alter the twist (*Tw*) and the writhe (*Wr*) of DNA. *Tw* computes the winding of each strand around the DNA axis, and *Wr* measures the non-planar turns of the DNA axis. Since Δ *Lk* = Δ *Tw* + Δ *Wr*, stabilization of +0.6 *Lk* units could result from over-twisting (Δ *Tw* > 0) and/or right-handed turning (Δ *Wr* > 0) of the CDEII and/or CDEIII segments. This scenario suggests that left-handed wrapping of DNA (Δ *Wr* < 0), as occurs in conventional nucleosomes, is not likely to happen at point centromeres. In such a case, CDEII and CDEIII would have to be severely over-twisted (e.g., DNA helical repeat of ~ 9 bp instead of 10.5 bp) in order to neutralize the negative *Wr* of DNA and constrain additional +0.6 *Lk* units. A more bearable situation would be that DNA does not turn or describes a flat U-turn at the point

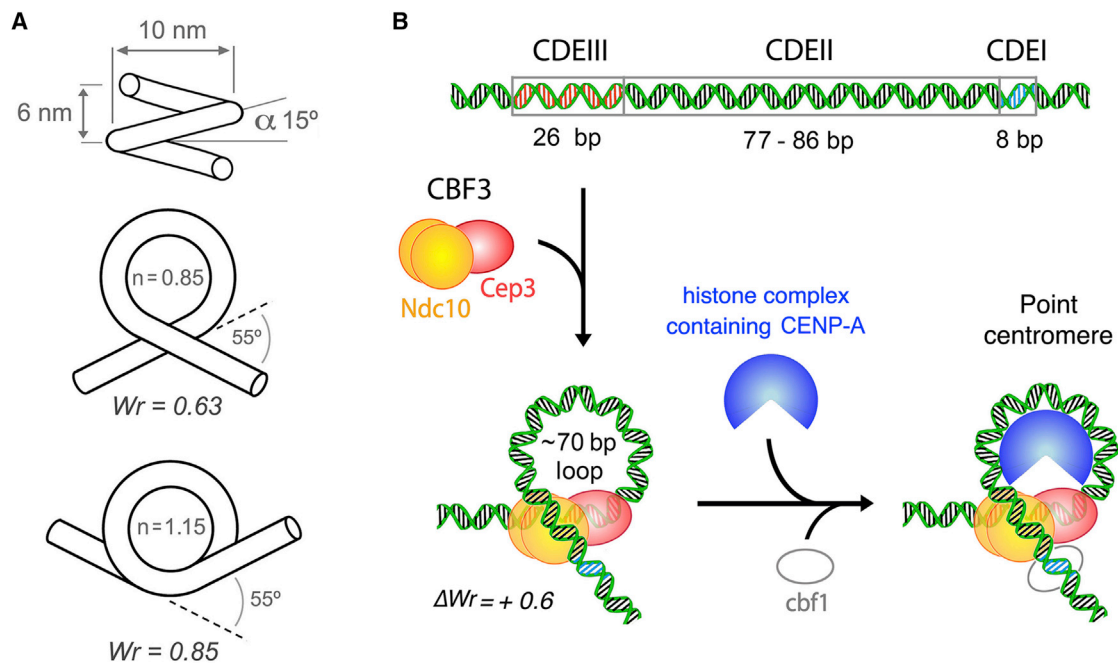


Figure 6. Model of DNA Topology and Global Architecture of the Yeast Point Centromere

(A) DNA writhe (Wr) of a right-handed coil that has pitch angle (α) of 15° and completes 0.85 or 1.15 helical turns (n). Angle α corresponds to a diameter/pitch ratio of 10/6 (diameter of nucleosomal DNA ~ 10 nm, DNA binding site separation in Ndc10 ~ 6 nm). Values of n (0.85 and 1.15) correspond to turns of $360^\circ \pm 55^\circ$. Wr is calculated following Fuller (1971) as $n(1 - \sin \alpha)$.

(B) Plausible assembly and architecture of a point centromere that constrains a ΔLk approximately $+0.6$. CDEIII is occupied by the CBF3 complex, in which the Cep3 homodimer (red) interacts with ~ 15 bp of the DNA, and one subunit of the Ndc10 homodimer (yellow) binds ~ 10 bp of DNA next to Cep3. For clarity, the Skp1-Ctf13 heterodimer of CBF3 is not shown. The second subunit of Ndc10 then interacts with a looped DNA segment in CDEII ~ 70 bp apart. This loop conforms a right-handed turn of dimensions $n = 0.85$ and $\alpha = 15^\circ$, such that its Wr is $+0.6$.

A histone complex that contains CENP-A is then recruited by CBF3 and accommodated in the looped DNA. The interaction of *cbf1* with CDEI stabilizes the full complex but does not alter its DNA topology.

centromere, such that its $Wr \sim 0$. In this case, DNA over-twisting across CDEII and CDEIII could suffice to constrain $+0.6 Lk$ units. However, the simplest explanation for the stabilization of $+0.6 Lk$ units is that DNA is wrapped in a right-handed manner at the point centromere ($\Delta Lk > 0$), as proposed in previous reports (Furuyama et al., 2006; Furuyama and Henikoff, 2009). This right-handed path could be determined by the CENP-A histone complex that is recruited by CBF3. However, there is no experimental evidence of inherently right-handed histone complexes. Moreover, (CENP-A/H4/H2B/H2A)₂ octamers assembled in vitro are left-handed (Sekulic et al., 2010; Tachiwana et al., 2011). This octamer configuration does not exclude that the horseshoe shape of sub-octameric particles containing CENP-A (tetrasomes or hemisomes) have the flexibility to accommodate DNA in a right- or left-handed path. This property is akin to (H3/H4)₂ tetrasomes, which can shift their chirality in function of the DNA supercoiling state (Hamiche et al., 1996) and by spontaneous fluctuations during nucleosome assembly (Vlijm et al., 2015). Then, in the case of the point centromere, the path of the DNA may be followed but not determined by histone-DNA interactions. Our results support a model in which a right-handed turn of DNA is determined by the CBF3 complex.

The assembly of the CBF3 complex is highly regulated and occurs prior to binding to DNA (Lechner and Carbon, 1991; Rus-

sell et al., 1999). CBF3 recognizes the CDEIII sequence via the Cep3 dimer, which contacts ~ 15 bp of CDEIII, including the essential CCG motif (Espelin et al., 1997; Pietrasanta et al., 1999; Russell et al., 1999; Purvis and Singleton, 2008). In contrast, the Ndc10 dimer of CBF3 does not recognize a specific DNA sequence. Ndc10 interacts with DNA inside and outside CDEIII, and also with CDEII, as it shows preferential binding to AT-rich regions (Espelin et al., 1997, 2003). These extended interactions initially suggested that several Ndc10 dimers concur in the point centromere. However, structural data revealed that the Ndc10 dimer binds two separate segments of DNA, each of ~ 10 bp (Cho and Harrison, 2012). Thus, while one subunit of Ndc10 binds DNA in defined register next to Cep3, the other subunit can interact with other DNA molecules or with a looped DNA segment located near CDEIII. Remarkably, this looping capacity was observed in earlier AFM studies of CBF3 bound to DNA (Pietrasanta et al., 1999). Specifically, the CBF3 complex shortened the DNA by ~ 70 bp and rendered an angle of 55° between the entry and exit segments of DNA. Thus, we calculated the theoretical Wr of a right-handed loop of 70 bp anchored by an Ndc10 dimer (Figure 6A). For a simple helix, $Wr = n(1 - \sin \alpha)$, n being the number of helical turns and α the angle of the helical pitch (Fuller, 1971). As the entry and exit DNA segments in the CBF3 complexes observed by AFM

formed an angle of 55° , we considered two plausible n values: $n = 1.15$ for a turn of $360^\circ + 55^\circ$, and $n = 0.85$ for a turn of $360^\circ - 55^\circ$. To determine the helical pitch, we considered that the superhelical turn was 6 nm high, which is the approximate DNA-binding site separation in Ndc10 (Cho and Harrison, 2012), and that the turn had a diameter of 10 nm, as for DNA wrapped on a histone core. This geometry produced $\alpha \sim 15^\circ$. Then, for $n = 1.15$, $Wr = 0.84$; and for $n = 0.85$, $Wr = 0.63$. Remarkably, the latter Wr value would account for the stabilization of $\Delta Lk +0.6$ without requiring significant alterations of Tw . Accordingly, we constructed a model of the point centromere in which the Ndc10 dimer configures a right-handed DNA turn ($n = 0.85$) by anchoring an entry segment at CDEIII (next to Cep3) and an exit segment at CDEII (Figure 6B). This configuration produces a loop of >70 bp, which could accommodate the CENP-A histone complex that is recruited via Scm3 (a CENP-A chaperone recognized by Ndc10) (Cho and Harrison, 2011; Wisniewski et al., 2014). Interestingly, the histone complex that interacts with CDEII elements *in vivo* has been found in variable bp registers (two orientations and two helical frames one turn apart) (Henikoff et al., 2014). This observation supports that this complex is accommodating into a predefined loop of DNA, rather than occupying a position determined by the DNA sequence. Our model is also consistent with the reported nuclease-protection data, which showed that point centromeres protect tightly a region of 120–135 bp (Cole et al., 2011; Krasovskiy et al., 2012). Finally, since CDEI is located outside the looped DNA, our model explains why CDEI and plausible interactions of *cbf1* with CBF3 are not relevant in determining the topology of DNA at the point centromere.

The singular topology of DNA constrained by point centromeres may not be an exclusive trait of yeast centromeric chromatin. The stabilization of positive supercoils has been observed *in vitro* when DNA is mixed with CENP-A/H4/H2B/H2A and the histone chaperone RbAp48 from *Drosophila* (Furuyama and Henikoff, 2009) and when chromatin is assembled with heterotypic histone-like CENP-S/T/W/X tetramers (Takeuchi et al., 2014). The partitioning locus of the yeast 2- μ m plasmid, which includes a CENP-A complex, induces also the stabilization of positive supercoils (Huang et al., 2011). Thus, positive supercoiling of DNA may be a general feature of centromeric chromatin, regardless of the rapid evolution of centromeric proteins (Malik and Henikoff, 2009). Left-handed wrapping of DNA in conventional nucleosomes confine DNA unwinding energy and allow specific structural transitions in response to the twisting and pulling forces generated by RNA and DNA polymerases. Centromeres undergo other mechanical processes. Right-handed wrapping of DNA in centromeric chromatin may serve to establish the bipolar orientation of the sister chromatids and to trigger distinctive structural transitions and checkpoint signals in response to the forces generated by the spindle.

EXPERIMENTAL PROCEDURES

Construction of Circular Minichromosomes

The yeast circular minichromosome TA1 (1,453 bp) and its derivatives of 1,686 bp and 1,286 bp were constructed as detailed in Supplemental Experimental Procedures. Oligonucleotides used to produce inserts are described

in Table S1, and detailed maps of the insert regions are illustrated in Figures S4 and S5. Linear fragments of each construct were circularized with T4 DNA ligase. Monomeric circles were gel-purified and used to transform the *S. cerevisiae* strain FY251.

Yeast Culture and DNA Extraction of Fixed Minichromosomes

Yeast colonies transformed with TA1-derived minichromosomes were grown at 26°C in standard yeast synthetic media containing TRP dropout supplement (Sigma) and 2% glucose. When the liquid cultures (20 ml) reached mid-log phase (optical density [OD] ~ 0.8), yeast cells were fixed *in vivo* by quickly mixing the cultures with one cold volume (-20°C) of ETol solution (95% ethanol, 28 mM toluene, 20 mM Tris HCl [pH 8.8], and 5 mM EDTA). As this fixation precludes the alteration of the *in vivo* Lk^{ch} values in subsequent manipulations (see Figure S2), the following steps were done at room temperature. Cells were sedimented, washed twice with water, resuspended in 400 μl TE, and transferred to a 1.5-ml microfuge tube containing 400 μl phenol and 400 μl acid-washed glass beads (425–600 μm , Sigma). Mechanic lysis of $>80\%$ cells was achieved by shaking the tubes in a FastPrep apparatus for 10 s at power 5. The aqueous phase of the lysed cell suspension was collected, extracted with chloroform, precipitated with ethanol, and resuspended in 100 μl Tris-EDTA buffer containing RNase-A. Following 15 min of incubation at 37° , the samples were extracted with phenol and chloroform, precipitated with ethanol, and resuspended in 30 μl of TE.

Solubilization of Native Minichromosomes

Yeast liquid cultures (20 ml) at mid-log phase (OD ~ 0.8) were not fixed with ETol solution. The cells were sedimented, washed with water, and resuspended in 500 μl cold (4°C) buffer L (10 mM Tris-HCl [pH 8], 1 mM EDTA, 1 mM EGTA, 1 mM DTT, 150 mM NaCl, Triton 0.1%, 1 mg/ml pepstatin, 1 mg/ml leupeptin, and 1 mM PMSF). The suspension was transferred to a 1.5-ml microfuge tube containing 500 μl acid-washed glass beads (425–600 μm , Sigma). Mechanic lysis of $>80\%$ cells was achieved after six cycles of 30 s of vortexing plus 30 s of ice cooling. The supernatant of the lysate was recovered by centrifugation ($2,000 \times g$ at 4°C) and loaded on a Sephacryl S-300 column equilibrated with buffer L. Circular minichromosomes eluted in the first filtration volume were adjusted to 8 mM MgCl_2 and 1mM ATP, pre-incubated at 26°C for 5 min, and then supplemented with control plasmids and catalytic amounts of purified topoisomerase I of vaccinia virus (Shuman et al., 1988) or *S. cerevisiae* topoisomerase II (Worland and Wang, 1989). Following incubations at 26°C for 10 min, reactions were quenched with one volume of 20 mM EDTA, 0.5% SDS, and 100 mg/ml proteinase K and incubated for 30 min at 60°C . The mixtures were extracted with phenol and chloroform, DNA precipitated with ethanol, and resuspended in 30 μl TE.

Indirect End-Labeling of the Micrococcal Nuclease-Digested Chromatin

Circular minichromosomes were solubilized and eluted from a Sephacryl S-300 column as described above and adjusted to 2 mM CaCl_2 . Following pre-incubation at 25°C for 5 min, micrococcal nuclease was added (2–100 U/ml) and digestions proceeded at 25°C for 5 min. Reactions were quenched with one volume of 20 mM EDTA, 0.5% SDS, 100 mg/ml proteinase K, and incubated for 60 min at 60°C . The mixtures were extracted with phenol and chloroform, DNA precipitated with ethanol and resuspended in 30 μl of TE. Digested DNA samples were singly cut with a restriction endonuclease, separated on an agarose gel, blotted, and probed with a short DNA sequence (<200 bp) contiguous to the single restriction site.

Electrophoresis of DNA Topoisomers

DNA extracted from yeast circular minichromosomes (samples of non-centromeric minichromosomes were diluted ~ 10 -fold) and the corresponding DNA circles relaxed *in vitro* with topoisomerase I at 26°C (temperature at which DNA topology was fixed *in vivo*) were loaded onto 1.2% (w/v) agarose gels. One-dimensional electrophoresis was carried out at 2.5 V/cm for 18 hr in Tris-borate EDTA (TBE) buffer (89 mM Tris-borate and 2 mM EDTA) containing 0.2 $\mu\text{g/ml}$ chloroquine. Two-dimensional electrophoresis was in TBE containing 0.1 or 0.2 $\mu\text{g/ml}$ chloroquine in the first dimension (2.5 V/cm for 18 hr) and in

TBE containing 1 $\mu\text{g/ml}$ chloroquine in the second dimension (5 V/cm for 4 hr). Gel-markers of *Lk* topoisomers were obtained by mixing partially relaxed samples of supercoiled DNA rings. Gels were blot-transferred to a nylon membrane and probed at 60°C with the 1,453-bp TRP1ARS1 sequence labeled with AlkPhos Direct (GE Healthcare). Chemiluminescent signals of increasing exposition periods were recorded on X-ray films.

DNA Topology Analyses

The midpoint of each *Lk* distribution, which does not necessarily coincide with the gel position of main topoisomers, was determined by quantifying with the ImageJ software the relative intensity of non-saturated signals of the individual *Lk* topoisomers. ΔLk was calculated as the distance (*Lk* units) between the midpoints of minichromosome (*Lk^{ch}*) and relaxed DNA (*Lk^r*) distributions, being both *Lk* distributions produced at the same temperature and resolved in the same gel electrophoresis.

SUPPLEMENTAL INFORMATION

Supplemental information includes Supplemental Experimental Procedures, five figures, and one table and can be found with this article online at <http://dx.doi.org/10.1016/j.celrep.2015.09.039>.

AUTHOR CONTRIBUTIONS

O.D.-I. and J.R. conceived the research, designed experiments, and analyzed data. O.D.-I., B.M.-G., J.S., and A.V. prepared materials and conducted experiments. J.R. wrote the manuscript.

ACKNOWLEDGMENTS

This research was supported by the Plan Nacional de I+D+I of Spain (grant BFU2011-23851 to J.R.).

Received: April 17, 2015

Revised: August 8, 2015

Accepted: September 14, 2015

Published: October 15, 2015

REFERENCES

- Aravamudhan, P., Felzer-Kim, I., and Joglekar, A.P. (2013). The budding yeast point centromere associates with two Cse4 molecules during mitosis. *Curr. Biol.* *23*, 770–774.
- Baker, R.E., Fitzgerald-Hayes, M., and O'Brien, T.C. (1989). Purification of the yeast centromere binding protein CP1 and a mutational analysis of its binding site. *J. Biol. Chem.* *264*, 10843–10850.
- Biggins, S. (2013). The composition, functions, and regulation of the budding yeast kinetochore. *Genetics* *194*, 817–846.
- Black, B.E., Jansen, L.E., Foltz, D.R., and Cleveland, D.W. (2010). Centromere identity, function, and epigenetic propagation across cell divisions. *Cold Spring Harb. Symp. Quant. Biol.* *75*, 403–418.
- Bloom, K.S. (2014). Centromeric heterochromatin: the primordial segregation machine. *Annu. Rev. Genet.* *48*, 457–484.
- Bloom, K.S., and Carbon, J. (1982). Yeast centromere DNA is in a unique and highly ordered structure in chromosomes and small circular minichromosomes. *Cell* *29*, 305–317.
- Bloom, K.S., Fitzgerald-Hayes, M., and Carbon, J. (1983). Structural analysis and sequence organization of yeast centromeres. *Cold Spring Harb. Symp. Quant. Biol.* *47*, 1175–1185.
- Bloom, K., Amaya, E., and Yeh, E. (1984). Centromeric DNA structure in yeast chromatin. In *Molecular Biology of the Cytoskeleton*, G.G. Borisy, D.W. Cleveland, and D.B. Murphy, eds. (Cold Spring Harbor Laboratory Press), pp. 175–184.
- Bui, M., Dimitriadis, E.K., Hoischen, C., An, E., Quénet, D., Giebe, S., Nita-Lazar, A., Diekmann, S., and Dalal, Y. (2012). Cell-cycle-dependent structural transitions in the human CENP-A nucleosome in vivo. *Cell* *150*, 317–326.
- Burrack, L.S., and Berman, J. (2012). Neocentromeres and epigenetically inherited features of centromeres. *Chromosome Res.* *20*, 607–619.
- Cai, M.J., and Davis, R.W. (1989). Purification of a yeast centromere-binding protein that is able to distinguish single base-pair mutations in its recognition site. *Mol. Cell. Biol.* *9*, 2544–2550.
- Camahort, R., Shivaraju, M., Mattingly, M., Li, B., Nakanishi, S., Zhu, D., Shilatifard, A., Workman, J.L., and Gerton, J.L. (2009). Cse4 is part of an octameric nucleosome in budding yeast. *Mol. Cell* *35*, 794–805.
- Cho, U.S., and Harrison, S.C. (2011). Recognition of the centromere-specific histone Cse4 by the chaperone Scm3. *Proc. Natl. Acad. Sci. USA* *108*, 9367–9371.
- Cho, U.S., and Harrison, S.C. (2012). Ndc10 is a platform for inner kinetochore assembly in budding yeast. *Nat. Struct. Mol. Biol.* *19*, 48–55.
- Clarke, L. (1998). Centromeres: proteins, protein complexes, and repeated domains at centromeres of simple eukaryotes. *Curr. Opin. Genet. Dev.* *8*, 212–218.
- Cole, H.A., Howard, B.H., and Clark, D.J. (2011). The centromeric nucleosome of budding yeast is perfectly positioned and covers the entire centromere. *Proc. Natl. Acad. Sci. USA* *108*, 12687–12692.
- Connelly, C., and Hieter, P. (1996). Budding yeast SKP1 encodes an evolutionarily conserved kinetochore protein required for cell cycle progression. *Cell* *86*, 275–285.
- Dalal, Y., Wang, H., Lindsay, S., and Henikoff, S. (2007). Tetrameric structure of centromeric nucleosomes in interphase Drosophila cells. *PLoS Biol.* *5*, e218.
- Depew, D.E., and Wang, J.C. (1975). Conformational fluctuations of DNA helix. *Proc. Natl. Acad. Sci. USA* *72*, 4275–4279.
- Dimitriadis, E.K., Weber, C., Gill, R.K., Diekmann, S., and Dalal, Y. (2010). Tetrameric organization of vertebrate centromeric nucleosomes. *Proc. Natl. Acad. Sci. USA* *107*, 20317–20322.
- Espelin, C.W., Kaplan, K.B., and Sorger, P.K. (1997). Probing the architecture of a simple kinetochore using DNA-protein crosslinking. *J. Cell Biol.* *139*, 1383–1396.
- Espelin, C.W., Simons, K.T., Harrison, S.C., and Sorger, P.K. (2003). Binding of the essential *Saccharomyces cerevisiae* kinetochore protein Ndc10p to CDEII. *Mol. Biol. Cell* *14*, 4557–4568.
- Fuller, F.B. (1971). The writhing number of a space curve. *Proc. Natl. Acad. Sci. USA* *68*, 815–819.
- Furuyama, S., and Biggins, S. (2007). Centromere identity is specified by a single centromeric nucleosome in budding yeast. *Proc. Natl. Acad. Sci. USA* *104*, 14706–14711.
- Furuyama, T., and Henikoff, S. (2009). Centromeric nucleosomes induce positive DNA supercoils. *Cell* *138*, 104–113.
- Furuyama, T., Dalal, Y., and Henikoff, S. (2006). Chaperone-mediated assembly of centromeric chromatin in vitro. *Proc. Natl. Acad. Sci. USA* *103*, 6172–6177.
- Furuyama, T., Codomo, C.A., and Henikoff, S. (2013). Reconstitution of hemisomes on budding yeast centromeric DNA. *Nucleic Acids Res.* *41*, 5769–5783.
- Gkikopoulos, T., Singh, V., Tsui, K., Awad, S., Renshaw, M.J., Scholfield, P., Barton, G.J., Nislow, C., Tanaka, T.U., and Owen-Hughes, T. (2011). The SWI/SNF complex acts to constrain distribution of the centromeric histone variant Cse4. *EMBO J.* *30*, 1919–1927.
- Haase, J., Mishra, P.K., Stephens, A., Haggerty, R., Quammen, C., Taylor, R.M., 2nd, Yeh, E., Basrai, M.A., and Bloom, K. (2013). A 3D map of the yeast kinetochore reveals the presence of core and accessory centromere-specific histone. *Curr. Biol.* *23*, 1939–1944.
- Hamiche, A., Carot, V., Alilat, M., De Lucia, F., O'Donohue, M.F., Revet, B., and Prunell, A. (1996). Interaction of the histone (H3-H4)₂ tetramer of the nucleosome with positively supercoiled DNA minicircles: Potential flipping of the

- protein from a left- to a right-handed superhelical form. *Proc. Natl. Acad. Sci. USA* **93**, 7588–7593.
- Hemmerich, P., Stoyan, T., Wieland, G., Koch, M., Lechner, J., and Diekmann, S. (2000). Interaction of yeast kinetochore proteins with centromere-protein/transcription factor Cbf1. *Proc. Natl. Acad. Sci. USA* **97**, 12583–12588.
- Henikoff, S., and Furuyama, T. (2010). Epigenetic inheritance of centromeres. *Cold Spring Harb. Symp. Quant. Biol.* **75**, 51–60.
- Henikoff, S., and Furuyama, T. (2012). The unconventional structure of centromeric nucleosomes. *Chromosoma* **121**, 341–352.
- Henikoff, S., Ramachandran, S., Krassovsky, K., Bryson, T.D., Codomo, C.A., Brogaard, K., Widom, J., Wang, J.P., and Henikoff, J.G. (2014). The budding yeast Centromere DNA Element II wraps a stable Cse4 hemisome in either orientation in vivo. *eLife* **3**, e01861.
- Hsu, J.M., Huang, J., Meluh, P.B., and Laurent, B.C. (2003). The yeast RSC chromatin-remodeling complex is required for kinetochore function in chromosome segregation. *Mol. Cell. Biol.* **23**, 3202–3215.
- Huang, C.C., Chang, K.M., Cui, H., and Jayaram, M. (2011). Histone H3-variant Cse4-induced positive DNA supercoiling in the yeast plasmid has implications for a plasmid origin of a chromosome centromere. *Proc. Natl. Acad. Sci. USA* **108**, 13671–13676.
- Jehn, B., Niedenthal, R., and Hegemann, J.H. (1991). In vivo analysis of the *Saccharomyces cerevisiae* centromere CDEIII sequence: requirements for mitotic chromosome segregation. *Mol. Cell. Biol.* **11**, 5212–5221.
- Jiang, C., and Pugh, B.F. (2009). A compiled and systematic reference map of nucleosome positions across the *Saccharomyces cerevisiae* genome. *Genome Biol.* **10**, R109.
- Jiang, W., Lechner, J., and Carbon, J. (1993). Isolation and characterization of a gene (CBF2) specifying a protein component of the budding yeast kinetochore. *J. Cell Biol.* **121**, 513–519.
- Krassovsky, K., Henikoff, J.G., and Henikoff, S. (2012). Tripartite organization of centromeric chromatin in budding yeast. *Proc. Natl. Acad. Sci. USA* **109**, 243–248.
- Lechner, J., and Carbon, J. (1991). A 240 kd multisubunit protein complex, CBF3, is a major component of the budding yeast centromere. *Cell* **64**, 717–725.
- Lochmann, B., and Ivanov, D. (2012). Histone H3 localizes to the centromeric DNA in budding yeast. *PLoS Genet.* **8**, e1002739.
- Malik, H.S., and Henikoff, S. (2009). Major evolutionary transitions in centromere complexity. *Cell* **138**, 1067–1082.
- Miell, M.D., Fuller, C.J., Guse, A., Barysz, H.M., Downes, A., Owen-Hughes, T., Rappsilber, J., Straight, A.F., and Allshire, R.C. (2013). CENP-A confers a reduction in height on octameric nucleosomes. *Nat. Struct. Mol. Biol.* **20**, 763–765.
- Mishra, P.K., Ottmann, A.R., and Basrai, M.A. (2013). Structural integrity of centromeric chromatin and faithful chromosome segregation requires Pat1. *Genetics* **195**, 369–379.
- Mizuguchi, G., Xiao, H., Wisniewski, J., Smith, M.M., and Wu, C. (2007). Nonhistone Scm3 and histones CenH3-H4 assemble the core of centromere-specific nucleosomes. *Cell* **129**, 1153–1164.
- Niedenthal, R.K., Sen-Gupta, M., Wilmen, A., and Hegemann, J.H. (1993). Cpf1 protein induced bending of yeast centromere DNA element I. *Nucleic Acids Res.* **21**, 4726–4733.
- Padeganeh, A., Ryan, J., Boisvert, J., Ladouceur, A.M., Dorn, J.F., and Maddox, P.S. (2013). Octameric CENP-A nucleosomes are present at human centromeres throughout the cell cycle. *Curr. Biol.* **23**, 764–769.
- Pederson, D.S., Venkatesan, M., Thoma, F., and Simpson, R.T. (1986). Isolation of an episomal yeast gene and replication origin as chromatin. *Proc. Natl. Acad. Sci. USA* **83**, 7206–7210.
- Pietrasanta, L.I., Thrower, D., Hsieh, W., Rao, S., Stemann, O., Lechner, J., Carbon, J., and Hansma, H. (1999). Probing the *Saccharomyces cerevisiae* centromeric DNA (CEN DNA)-binding factor 3 (CBF3) kinetochore complex by using atomic force microscopy. *Proc. Natl. Acad. Sci. USA* **96**, 3757–3762.
- Prunell, A. (1998). A topological approach to nucleosome structure and dynamics: the linking number paradox and other issues. *Biophys. J.* **74**, 2531–2544.
- Purvis, A., and Singleton, M.R. (2008). Insights into kinetochore-DNA interactions from the structure of Cep3Delta. *EMBO Rep.* **9**, 56–62.
- Rhee, H.S., Bataille, A.R., Zhang, L., and Pugh, B.F. (2014). Subnucleosomal structures and nucleosome asymmetry across a genome. *Cell* **159**, 1377–1388.
- Russell, I.D., Grancell, A.S., and Sorger, P.K. (1999). The unstable F-box protein p58-Ctf13 forms the structural core of the CBF3 kinetochore complex. *J. Cell Biol.* **145**, 933–950.
- Schueler, M.G., and Sullivan, B.A. (2006). Structural and functional dynamics of human centromeric chromatin. *Annu. Rev. Genomics Hum. Genet.* **7**, 301–313.
- Segal, E., Fondufe-Mittendorf, Y., Chen, L., Thåström, A., Field, Y., Moore, I.K., Wang, J.P., and Widom, J. (2006). A genomic code for nucleosome positioning. *Nature* **442**, 772–778.
- Sekulic, N., Bassett, E.A., Rogers, D.J., and Black, B.E. (2010). The structure of (CENP-A-H4)₂ reveals physical features that mark centromeres. *Nature* **467**, 347–351.
- Shuman, S., Golder, M., and Moss, B. (1988). Characterization of vaccinia virus DNA topoisomerase I expressed in *Escherichia coli*. *J. Biol. Chem.* **263**, 16401–16407.
- Tachiwana, H., Kagawa, W., Shiga, T., Osakabe, A., Miya, Y., Saito, K., Hayashi-Takanaka, Y., Oda, T., Sato, M., Park, S.Y., et al. (2011). Crystal structure of the human centromeric nucleosome containing CENP-A. *Nature* **476**, 232–235.
- Takeuchi, K., Nishino, T., Mayanagi, K., Horikoshi, N., Osakabe, A., Tachiwana, H., Hori, T., Kurumizaka, H., and Fukagawa, T. (2014). The centromeric nucleosome-like CENP-T-W-S-X complex induces positive supercoils into DNA. *Nucleic Acids Res.* **42**, 1644–1655.
- Thoma, F., Bergman, L.W., and Simpson, R.T. (1984). Nuclease digestion of circular TRP1ARS1 chromatin reveals positioned nucleosomes separated by nuclease-sensitive regions. *J. Mol. Biol.* **177**, 715–733.
- Vlijm, R., Lee, M., Lipfert, J., Lusser, A., Dekker, C., and Dekker, N.H. (2015). Nucleosome assembly dynamics involve spontaneous fluctuations in the handedness of tetrasomes. *Cell Rep.* **10**, 216–225.
- Wilmen, A., Pick, H., Niedenthal, R.K., Sen-Gupta, M., and Hegemann, J.H. (1994). The yeast centromere CDEI/Cpf1 complex: differences between in vitro binding and in vivo function. *Nucleic Acids Res.* **22**, 2791–2800.
- Wisniewski, J., Hajj, B., Chen, J., Mizuguchi, G., Xiao, H., Wei, D., Dahan, M., and Wu, C. (2014). Imaging the fate of histone Cse4 reveals de novo replacement in S phase and subsequent stable residence at centromeres. *eLife* **3**, e02203.
- Worland, S.T., and Wang, J.C. (1989). Inducible overexpression, purification, and active site mapping of DNA topoisomerase II from the yeast *Saccharomyces cerevisiae*. *J. Biol. Chem.* **264**, 4412–4416.
- Xiao, H., Mizuguchi, G., Wisniewski, J., Huang, Y., Wei, D., and Wu, C. (2011). Nonhistone Scm3 binds to AT-rich DNA to organize atypical centromeric nucleosome of budding yeast. *Mol. Cell* **43**, 369–380.
- Zhang, W., Colmenares, S.U., and Karpen, G.H. (2012). Assembly of *Drosophila* centromeric nucleosomes requires CID dimerization. *Mol. Cell* **45**, 263–269.
- Zlatanova, J., Bishop, T.C., Victor, J.M., Jackson, V., and van Holde, K. (2009). The nucleosome family: dynamic and growing. *Structure* **17**, 160–171.

Cell Reports

Supplemental Information

DNA Topology and Global Architecture of Point Centromeres

**Ofelia Díaz-Ingelmo, Belén Martínez-García, Joana Segura, Antonio Valdés, and
Joaquim Roca**

Supplemental Experimental Procedures

Detailed construction of yeast circular minichromosomes

To construct the minichromosome TA1, the yeast TRP1ARS1 *EcoRI-EcoRI* genomic fragment (1453 bp) was cloned in the plasmid pSPT18 (boehringer mannheim®). Circular DNA molecules were obtained by self-circularization of the excised 1453-bp segment. Monomeric circles were gel purified and electroporated into yeast cells. Individual transformants were selected on TRP dropout plates.

To construct the 1686-bp circular minichromosomes, the *EcoRI* site of pSPT18 that flanked the TRP1 promoter was substituted by a *Sall* site (primers 0F/0R) (Table S1). In the resulting construct, unique *BamHI* and *KpnI* sites upstream *Sall* were used to insert new elements between the nucleosomes I and nucleosome VII of TA1. *CEN4* was inserted as a 260 bp *BamHI-KpnI* fragment obtained from the yeast plasmid Ycp50. From the above construct, *mutCEN4* was generated by disrupting the CDEIII sequence of *CEN4*. To do this, the essential motif of CDEIII (Jenh et al. 1991) was changed from CCGAAA to CATATG via PCR (primers 1F/1R). The DNA sequence of yeast nucleosome *High2* was inserted as a *BamHI-KpnI* 260 bp fragment obtained via PCR (primers 2F/2R). After sequence verification, the TA1 segments (containing the *CEN4*, *CEN4mut*, or *High2* inserts) were excised from pSPT18 as *Sall-Sall* fragments of 1686 bp. These DNA molecules were circularized and the monomeric rings purified to transform yeast. Figure S4 shows the position of the inserts between nucleosomes I and VII of TA1 in the resulting 1686-bp minichromosomes.

To construct the 1286-bp circular minichromosomes, PCR was used to generate a pair of *NotI* sites in TA1 (primers 3F/3R), which allowed the circularization of a 1040-bp segment that contained the TRP1ARS1 sequences but eliminated the nucleosomes VI and VII of TA1. DNA segments of 246 bp were then generated via PCR and inserted in the unique *BamHI-KpnI* site located between nucleosomes I and V of TA1. These segments contained: *CEN4* (primers 5F/5R), *mutCEN* (primers 1F/1R), *High2* (primers 6F/6R), *CEN2* (primers 8F/8R), *CEN7* (primers 10F/10R), *CEN12* (primers 11F y 11R) and *CEN4* in reverse direction (primers 12F/12R). From the *CEN4* construct, Δ *CDEI* was generated by disrupting CDEI. For this purpose, the central motif of CDEI (Wilmen et al, 1994) was changed from GTCACATG to CATCGATC via PCR (primers 4F/4R). From the *CEN4* and Δ *CDEI* constructs, Δ *CDEII* was generated by engineering *NcoI* and *XhoI* sites at the edges of CDEII and then replacing the CDEII segment with 76 bp of the *High2* sequence (Table S1). After sequence verification, the TRP1ARS1 sequences containing the above-mentioned inserts were excised from pSPT18 as *NotI-NotI* fragments of 1286 bp. These DNA fragments were circularized and the monomeric rings purified to transform yeast. Figure S5 shows the position of the inserts between nucleosomes I and V of TA1 in the resulting 1286-bp minichromosomes.

Table S1. Oligonucleotides used to produce Inserts. Related to Experimental Procedures.

Primer pairs (5' to 3') for the production of specific inserts via PCR

0F CGTGG AATTCCATATGGTCGACTATTGAGCACCGTGAGTATACGTG
0R GACTAGTCCATATGTGCGACTATTGAGCACGTGAGTATACG

1F GATTACATATGCATAAAAACCTGCTCAAGAAAAAGAACTG
1R GGT TTTATGCATATGTAATCATAAACATAAATAAATTAT TTC

2F CGCGGATCCTATGAACGGTAAGCAGGGAG
2R CGGGGTACCTTTCATGGCCATCTTTGTTCC

3F CGCGGAATTCGGTCGAAAAAGAAAAGG-
3R CAAGTGCGGCCGCCATTCTCTCAACAAGTTTGATTCC

4F AGGATCGATGCTTTTGCAAGCAATTAATTTTGC
4R AAGCATCGATCCTTATAATCAACTTTTTTAAAAATTTAAAATAC

5F GGCGGATCCACTTTTGCCTAAATCACA AATTGCAAATTTAATTGCTTG
5R GCCGGTACCTAACATGTGGA ACTCGAGCTCGGATCGATTTTG GCC

6F CGGGGATCCTATGAACGGTAAGCAGGGAGTTAAC
6R CGGGGTACCTCTTTGTTCC TTTCACTATTTTGC

8F CGGGGTACCTAATAATAAATTAATCTTGAGCAAATTG
8R GCGGGATCCCATACATGAAGAATATAGGATCTG

9F CGGGGTACCAACTTTTGCCTAAATCACA AATTGC
9R GCGGGATCCTATAATGCCTATTTTCTGTAAAGAG

10F CGCGGATCCCCTTTCTTG CATTATAATTATCC
10R CGCGGTACCAACATCAATATATATAATATAATCACAT TTC

11F CGCGGATCCCAACGAGTAAACGCTTACTTTTCTAAAAG
11R CGCGGTACCAAGAAAGAGAATGATGAAAAA A A A A A CAGG

12F GCGGGATCCCAACTTTTGCCTAAATCACA AATTGC
12R GCGGGATCCTATAATGCCTATTTTCTGTAAAGAG

DNA sequence of *High2* used to replace CDEII

CCATGGCTCCAGCTTCGCGTTCGTGAGAGCGACGTCTTTTCCCCGCATCGCCACCGTACAATTCATTAA**ACTCGAG**

(added *NcoI* and *XhoI* sites are indicated in bold)

Supplemental Figures

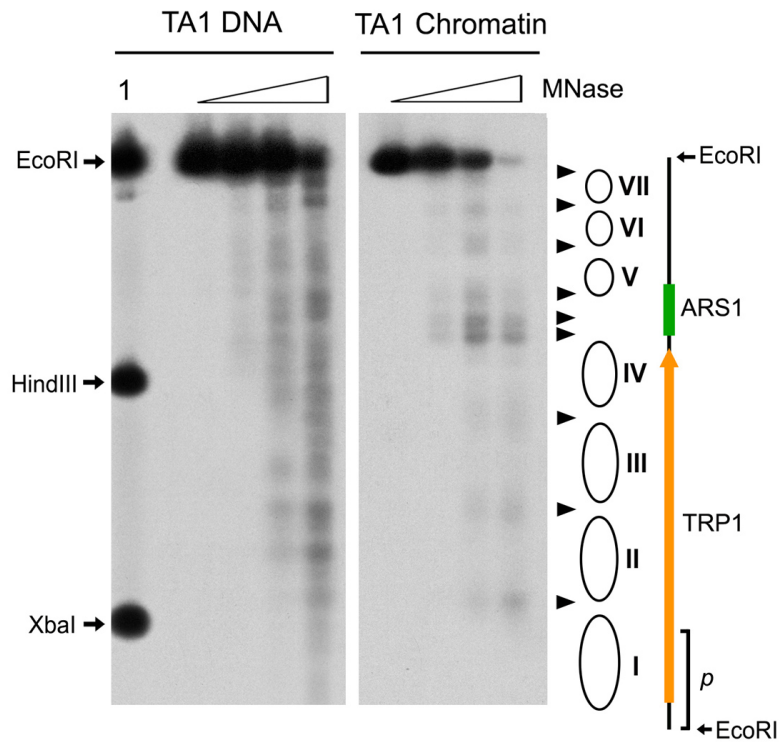


Figure S1. DNA cleavage sites by micrococcal nuclease on the TA1 minichromosome. Related to Figure 1A. Yeast cells containing the TA1 circular minichromosome were cultured to mid-log phase. The native minichromosome was solubilized and digested with increasing amounts of micrococcal nuclease (MNase) as described in the methods section. Similar digestions were performed on naked TA1 DNA rings. Digested DNA samples were restricted with endonuclease *EcoRI*, separated on a 1% agarose gel, blotted and probed with the 186 bp *EcoRI-XbaI* TRP1 fragment (*p*). Lane 1, *XbaI* and *HindIII* site markers. Arrowheads indicate preferential cleavage sites at the TA1 minichromosome and their correspondence with nucleosomal linker regions including the hypersensitive ARS1 region. The seven nucleosome positions (I-VII) inferred from this experiment coincide with those previously determined by Thoma *et al.* (1984).

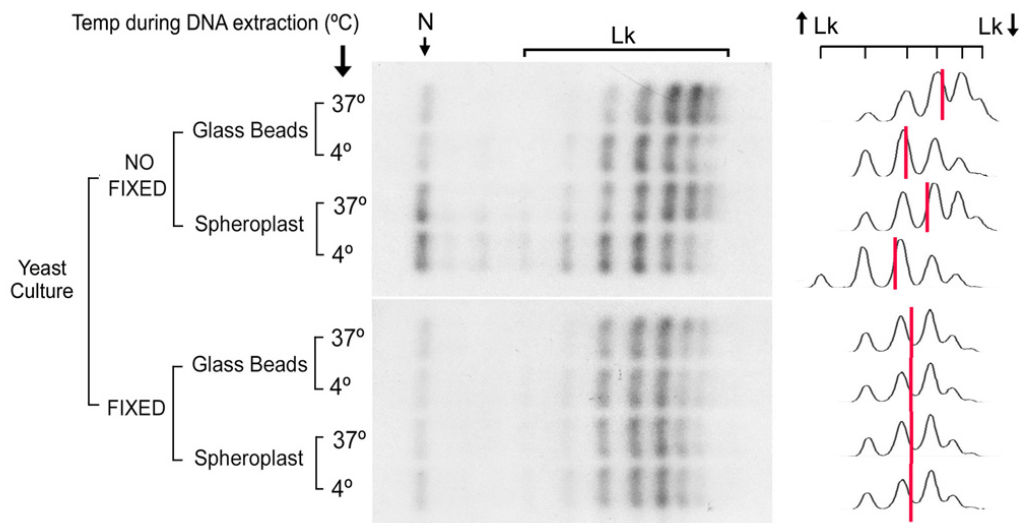


Figure S2. Effect of temperature and DNA extraction procedure on the *Lk* of yeast circular minichromosomes. Related to Experimental Procedures.

Yeast cells containing the TA1 minichromosome (1453 bp) were cultivated in liquid media (100 ml) at 26°C. At mid-log phase, one half of the culture was quenched by quickly mixing it with one volume (50 ml) of cold (-20°C) ETol solution (Ethanol 95%, Toluene 28 mM, Tris HCl pH 8.8 20 mM, EDTA 5 mM). The fixed and non-fixed cultures were then processed in parallel to extract yeast DNA by different procedures. One half of each fraction was used to produce spheroplast by standard procedures. Lysis of the spheroplasts prior to DNA extraction was done at 37°C in one half of the sample and at 4°C in the other half. The remaining cells of the fixed and non-fixed cultures were lysed and their DNA extracted by shaking with glass-beads in presence of phenol. The lysis was done at 37°C in one half of the sample and at 4°C in the other half. The 8 samples of DNA obtained were loaded onto 1.2% agarose gels and electrophoresed at 2.5 V/cm for 18 h in 1xTBE buffer containing 0.2 µg/ml chloroquine. The gel-blot was probed with TA1. The position of nicked circles (N) and linking number topoisomers (*Lk*) is indicated. Plots and intensity centers of *Lk* distributions were obtained with ImageJ. The results show that, in the non-fixed cultures, *Lk* distributions markedly diverged depending on the DNA extraction procedure and temperature. *Lk* values were higher at 4°C than at 37°C since the helical repeat of DNA lessens as the temperature drops (Depew and Wang 1975). Therefore, endogenous topoisomerases quickly adjusted the *Lk* of DNA prior cell lysis. Conversely, in the cultures fixed at 26°C, the *Lk* distributions were nearly identical irrespective of the DNA extraction conditions. Therefore, *in vivo* fixation inactivates the cellular topoisomerases so preventing *Lk* variations due to temperature or chromatin disruption during subsequent manipulation of the samples.

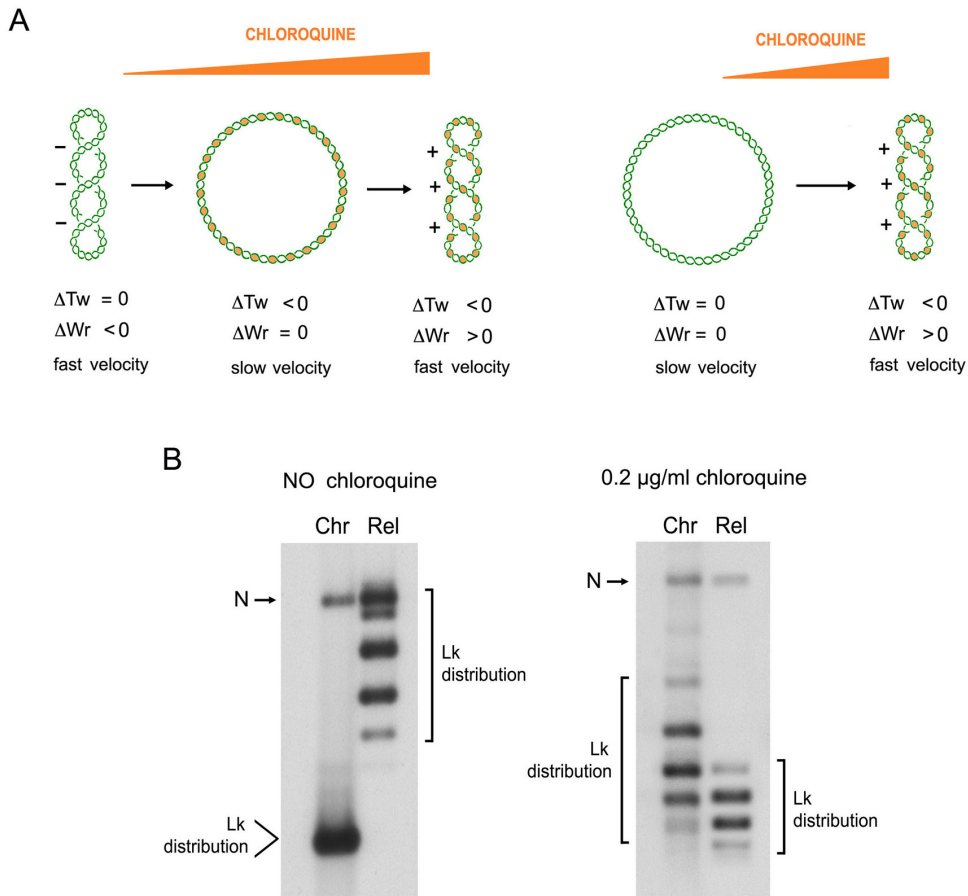


Figure S3. Effect of chloroquine on the gel velocity of *Lk* topoisomers. Related to Figure 1B. A) Chloroquine intercalation unwinds the double helix and so reduces the twist (T_w) of the DNA. In a covalently closed DNA circle of a given linking number (L_k), this change of T_w is compensated with a change of W_r (DNA writhe or number of supercoil crossings) since $\Delta L_k = \Delta T_w + \Delta W_r$. The scheme illustrates the effects of chloroquine on the T_w and W_r of DNA circles that are negatively supercoiled (left) and relaxed (right). These DNA deformations modify the gel velocity of individual L_k topoisomers, which depends on their number of supercoils (absolute W_r). B) In the electrophoresis with no chloroquine, DNA from TA1 minichromosomes (Chr) is negatively supercoiled. All L_k topoisomers have $W_r < 0$ and the L_k distribution is collapsed in a fast migration band. Relaxed TA1 DNA (Rel) produces a L_k distribution with some topoisomers overlapping with the nicked form (N) and others migrating slightly faster ($W_r > 0$). By running the same DNA samples in presence of 0.2 $\mu\text{g/ml}$ of chloroquine, the minichromosome DNA (Chr) reduces its absolute W_r such that all the L_k topoisomers are resolved as individual bands. At the same time, the L_k distribution of relaxed DNA (Rel) increases its absolute W_r , individual topoisomers migrate faster and no longer overlap with the nicked form (N).

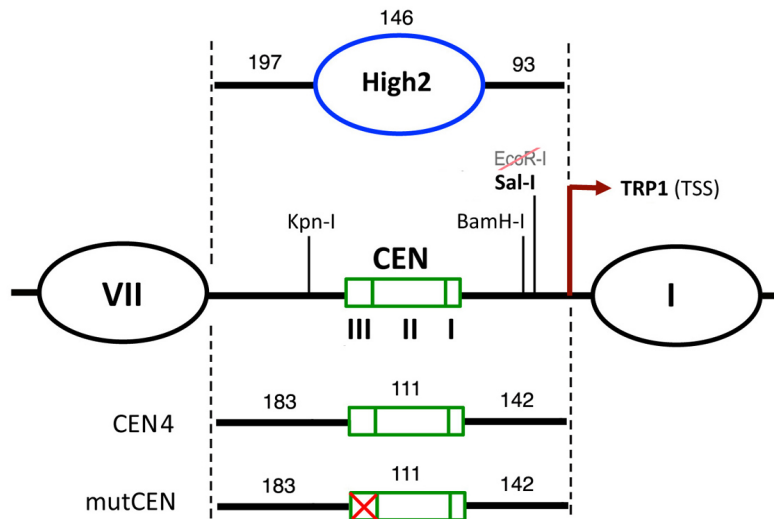


Figure S4. Detailed maps of the elements inserted between nucleosomes I and VII of TA1 in 1686-bp minichromosomes. Related to Figure 2A. *KpnI-BamHI* sites used to insert the *High2*, *CEN4* and *mutCEN* sequences in between nucleosomes I and VII of TA1 are indicated. *EcoRI* was replaced by *SalI* as the circularization site of the minichromosome. III, II, I denotes the CDEs position. The length (bp) of the inserted elements and their linker distance to the edge of nucleosome VII and to the transcription start site (TSS) of TRP1 are indicated.

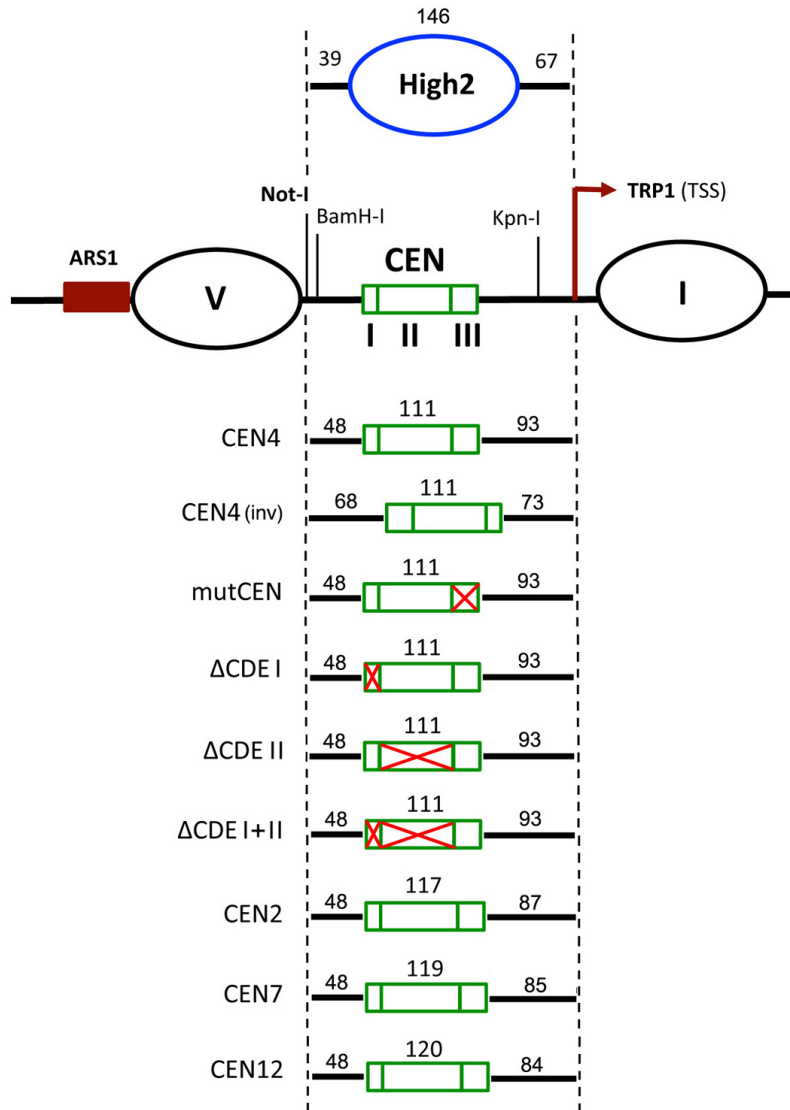


Figure S5. Detailed maps of the elements inserted between nucleosomes I and V of TA1 in 1286-bp minichromosomes. Related to Figures 3A, 4B and 5A. *KpnI-BamHI* sites used to insert the *High2* and distinct CEN sequences in between nucleosomes I and V of TA1 are indicated. *NotI* is the circularization site of the minichromosome. III, II, I denote the CDEs positions. The length (bp) of the inserted elements and their linker distance to the edge of nucleosome V (*NotI* site) and to the transcription start site (TSS) of TRP1 are indicated.

artículo #2

RESUMEN ARTÍCULO #2

“Intracellular nucleosomes constrain a DNA linking number difference of -1.26 that reconciles the Lk paradox”

Joana Segura, Ricky S. Joshi, Ofelia Díaz-Ingelmo, Antonio Valdés, Silvia Dyson, Belén Martínez-García & Joaquim Roca.






La interrelación entre la estructura de la cromatina y la topología del ADN juega un papel fundamental para la regulación de la actividad génica y la arquitectura de los cromosomas. Aun así, la topología local del ADN a lo largo de los elementos de la cromatina ha sido muy poco investigada. Ello es debido, en parte, a su complejidad conceptual y a su dificultad metodológica. Un caso paradigmático es la llamada “paradoja del número de enlace del ADN nucleosomal”, que ha sido objeto de debate durante décadas. Esta paradoja se refiere a la incongruencia entre el valor esperado de diferencia del número de enlace (ΔLk) estabilizado por cada nucleosoma ($\Delta Lk \approx -2$) y el valor estimado experimentalmente de $\Delta Lk \approx -1$. En este trabajo, hemos abordado este problema analizando la topología del ADN de una librería de mononucleosomas insertados en pequeños minicromosomas, de manera que nos permite determinar con precisión el valor medio de ΔLk estabilizado por nucleosomas individuales *in vivo*. Nuestros resultados indican que la mayoría de nucleosomas estabilizan alrededor de -1.26 unidades de ΔLk . Teniendo en cuenta que las deformaciones del ADN nucleosomal en forma de ΔTw y ΔWr ($\Delta Tw \approx +0.2$; $\Delta Wr \approx -1.5$) predicen un valor de ΔLk de -1.3 ($\Delta Lk = \Delta Tw + \Delta Wr$), nuestros resultados resuelven la discrepancia histórica entre el valor teórico y el experimental.

ARTICLE

DOI: 10.1038/s41467-018-06547-w

OPEN

Intracellular nucleosomes constrain a DNA linking number difference of -1.26 that reconciles the Lk paradox

Joana Segura ¹, Ricky S. Joshi¹, Ofelia Díaz-Ingelmo ¹, Antonio Valdés ¹, Silvia Dyson ¹, Belén Martínez-García¹ & Joaquim Roca ¹

The interplay between chromatin structure and DNA topology is a fundamental, yet elusive, regulator of genome activities. A paradigmatic case is the “linking number paradox” of nucleosomal DNA, which refers to the incongruence between the near two left-handed superhelical turns of DNA around the histone octamer and the DNA linking number difference (ΔLk) stabilized by individual nucleosomes, which has been experimentally estimated to be about -1.0 . Here, we analyze the DNA topology of a library of mononucleosomes inserted into small circular minichromosomes to determine the average ΔLk restrained by individual nucleosomes in vivo. Our results indicate that most nucleosomes stabilize about -1.26 units of ΔLk . This value balances the twist ($\Delta Tw \approx +0.2$) and writhe ($\Delta Wr \approx -1.5$) deformations of nucleosomal DNA in terms of the equation $\Delta Lk = \Delta Tw + \Delta Wr$. Our finding reconciles the existing discrepancy between theoretical and observed measurement of the ΔLk constrained by nucleosomes.

¹Molecular Biology Institute of Barcelona (IBMB), Spanish National Research Council (CSIC), Barcelona 08028, Spain. Correspondence and requests for materials should be addressed to J.R. (email: joaquim.roca@ibmb.csic.es)

Cellular DNA is packaged into chromatin via a hierarchical series of folding steps. The basic packaging unit, the nucleosome, contains about 147 base pairs (bp) of core DNA, making near two left-handed superhelical turns around a histone octamer¹. However, nucleosomes are not uniform and static entities. They can present positional instability, conformational fluctuations, histone variants, and histone modifications^{2–4}, all of which play a major role in the regulation of chromatin architecture and genome transactions^{5–7}. However, some fundamental aspects of nucleosomes, such as their interplay with DNA topology, remain elusive. In this respect, a paradigmatic case is the so-called “linking number paradox” of nucleosomal DNA^{8–10}, which has been the subject of debate for decades^{11,12}.

The linking number (Lk) of DNA is the number of times the single strands of the duplex intertwine around each other^{12,13}. The Lk paradox refers to the discrepancy between the theoretical and the experimental measurement of the DNA linking number difference (ΔLk) stabilized by nucleosomes. According to the general equation $\Delta Lk = \Delta Tw + \Delta Wr$ ¹⁴, it was expected that a nucleosome should stabilize a ΔLk value close to -2 , considering that DNA describes near two left-handed superhelical turns ($\Delta Wr \approx -2$) and assuming no significant changes in the double helical DNA twist ($\Delta Tw \approx 0$). However, numerous studies have persistently concluded that the ΔLk constrained by individual nucleosomes is ~ -1.0 . In those experiments, circular DNA molecules with and without nucleosomes were relaxed with a topoisomerase, and ΔLk was calculated. Most of these experiments used the simian virus 40 (SV40) minichromosome as a chromatin model. SV40 was found to have a ΔLk of about -26 ^{15,16}, which was comparable to the number of nucleosomes (24 to 27) typically observed by electron microscopy^{17,18}. This ΔLk value, which applied to the histone H1-containing native minichromosome, also held true for the H1-free SV40 minichromosome reconstituted in vitro from naked DNA and the four core histones¹⁹. Another study performed with the yeast circular minichromosome TRP1ARS1 harboring seven nucleosomes also concluded a ΔLk value of -1 per nucleosome²⁰. Finally, in vitro experiments of chromatin reconstitution using tandem repeats of nucleosome positioning sequences and core histones indicated ΔLk values of -1.01 ± 0.08 ²¹ and -1.04 ± 0.08 ²² per nucleosome.

The first hypothesis put forward to explain the Lk paradox was that core DNA was notably overtwisted ($\Delta Tw \approx +0.7$)^{9,23}, which meant that the helical periodicity (h) of DNA was smaller in the nucleosome than in free DNA in the solution ($h \approx 10.5$ bp/turn)²⁴. The plausible overtwisting of nucleosomal DNA was then calculated by $\Delta Tw = \Delta \emptyset + \Delta STw$, in which the winding number (\emptyset) depends on the helical repeat of DNA at the nucleosome surface (hs), and the surface twist (STw) is a correction function that accounts for the curved path of DNA in the nucleosome²⁵. Multiple measurements of hs based on DNase I footprinting^{23,26–28}, hydroxyl radical cleavage^{29,30}, and DNA base-pair periodicity^{31–33} indicated that the mean value of hs is about 10.2 bp/turn, which implied that $\Delta \emptyset \approx +0.4$. The nucleosomal STw was calculated from a derivation for a straight solenoidal helix to be -0.19 ^{34,35}. These figures indicated that the overall ΔTw of the core DNA is about $+0.2$, a value that was later corroborated by its direct measurement on the nucleosome structure at atomic resolution¹. The structural data showed also that the core DNA describes about 1.65 left-handed superhelical turns with a pitch angle of about 4 degrees, which produce a ΔWr value of about -1.5 ^{11,36}. The Tw and Wr deformations of the core DNA were therefore not sufficient to explain the Lk paradox.

A second hypothesis to explain the paradox pointed to the topology of DNA outside the core region. The zig-zag architecture observed in some nucleosomal fibers led to the proposal

that if linker DNA segments were repeatedly crossed with a similar geometry, the overall writhe (ΔWr) of the nucleosomal fiber would increase and produce the apparent $\Delta Lk \approx -1$ per nucleosome³⁷. However, recent modeling and experimental measurements with regular arrays of positioned nucleosomes demonstrated that their ΔLk varies markedly with nucleosome spacing, such that the apparent ΔLk value per nucleosome can range from -0.8 to -1.4 depending on the DNA linker length³⁸. Another proposal involving the topology of DNA outside the core region was based on the study of single nucleosomes reconstituted on small DNA circles³⁹ and on the torsional resilience of nucleosomal fibers in vitro⁴⁰. These studies suggested that nucleosomes fluctuate between three conformations: one in which incoming and outgoing linker segments form a negative crossing, one with uncrossed linkers, and one in which the linker segments cross positively. As a result, the average ΔWr of nucleosomal DNA would be reduced, as would its ΔLk . However, since these fluctuations depend on external constraints and forces, their plausible relevance to explain the Lk paradox is uncertain.

Here we revisit the Lk paradox of nucleosomal DNA by measuring the ΔLk constrained by individual nucleosomes in intracellular chromatin. As a chromatin model, we use small circular minichromosomes of budding yeast, whose nucleosomes are structurally identical to that of higher eukaryotes⁴¹ and are mainly depleted of linker histones⁴². First, we determine the ΔLk constrained in minichromosomes containing a known number of nucleosomes. To this end, we fix their DNA topology in vivo and compare it with that of naked DNA relaxed in vitro. We then insert a library of mononucleosomes into these minichromosomes and calculate the Lk gain ($\Delta \Delta Lk$) produced by the individual nucleosomes. Our results indicate that the majority of nucleosomes stabilize -1.26 units of ΔLk . We conclude that this experimental ΔLk value, along with the calculated twist and writhe deformations of DNA upon nucleosome formation, provides a solution for the Lk paradox.

Results

The ΔLk constrained by circular minichromosomes in vivo. For the purpose of the present study, we constructed YCp1.3, a circular minichromosome of *S. cerevisiae* comprising only 1341 bp. In order to ensure stable replication and segregation, YCp1.3 contained the genomic *TRP1-ARS1* segment and the point centromere *CEN2* of yeast (Fig. 1a). The *TRP1-ARS1* segment has four nucleosomes (I to IV) occupying the *TRP1* coding sequence and a fifth nucleosome (V) positioned 5' from the *ARS1* element^{43,44}. *CEN2* was then allocated between nucleosomes V and I, upstream of the *TRP1* promoter (see Supplementary Fig. 1 for detailed configuration and the bp sequence of YCp1.3). We confirmed the chromatin organization of YCp1.3 via micrococcal nuclease digestion of the minichromosome solubilized from yeast cells. As expected, the DNA sites most sensitive to nuclease digestion occurred in the *ARS1* region, followed by the sites corresponding to the linker DNA regions of the point centromere and the five nucleosomes (Fig. 1b).

Next, we determined the ΔLk constrained by the chromatin structure of YCp1.3. This value is the difference between the distribution of Lk values of the minichromosome in vivo (Lk^{CHR}) and that of the relaxed DNA in vitro (Lk^0). To this end, we fixed the Lk values of the minichromosome in vivo by quenching a culture of yeast cells with a freezing ethanol–toluene solution. We showed in previous studies that this quick fixation step irreversibly inactivates cellular topoisomerases, thereby precluding plausible alterations of Lk^{CHR} during cell disruption and DNA extraction⁴⁵. Since Lk^0 depends on temperature⁴⁶, we relaxed the naked YCp1.3 DNA circle with a type-1B

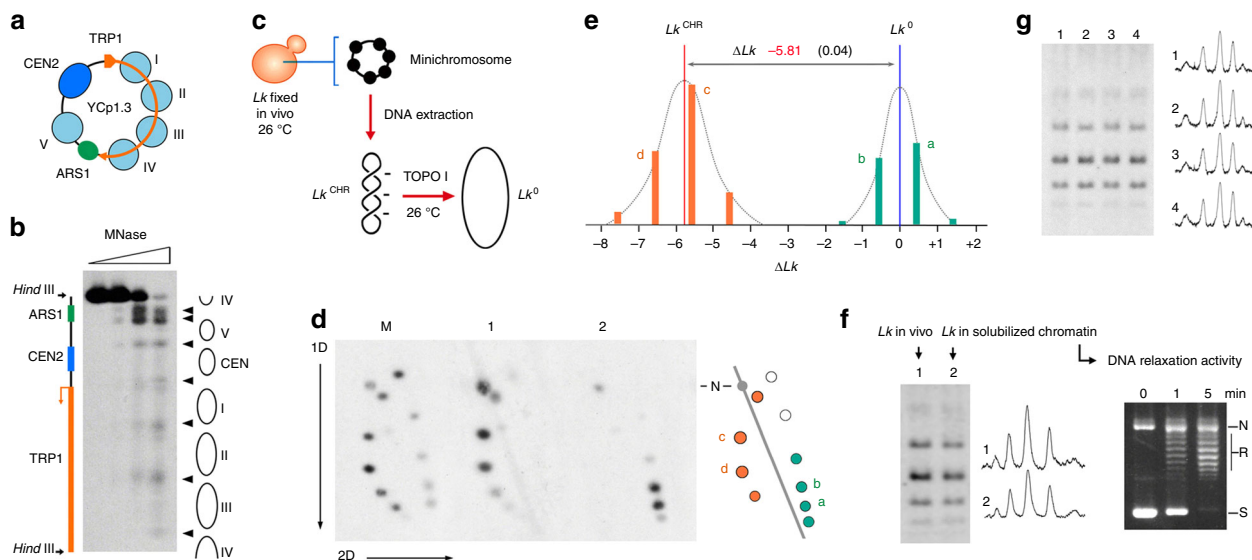


Fig. 1 Structure and DNA linking number difference in the yeast YCp1.3 minichromosome. **a** Scheme of YCp1.3 (1341 bp) indicating the position of the five nucleosomes (I to V) that occupy the genomic *TRP1-ARS1* segment of *S. cerevisiae*. **b** Micrococcal nuclease digestion pattern of YCp1.3. Nuclease-sensitive sites (arrow heads) are indicated and aligned to functional and structural elements of YCp1.3. **c** Experimental setting to obtain the *Lk* distribution of the YCp1.3 minichromosome in vivo (Lk^{CHR}) and the *Lk* distribution of its DNA relaxed in vitro (Lk^0). **d** Two-dimensional (2D) electrophoresis of the DNA of the YCp1.3 minichromosome extracted from cells fixed at 26 °C (lane 1) and following relaxation of the naked DNA with topoisomerase I at 26 °C (lane 2). A marker of *Lk* topoisomers of YCp1.3, in which *Lk* values increase clockwise, is included (lane M). DNA electrophoresis, blotting and probing were done as described in the methods. The 2D scheme (right) depicts the relative position of *Lk* topoisomers visible in lane 1 (orange dots) and lane 2 (green dots). Most intense *Lk* topoisomers (a, b, c, d) and nicked (N) molecules are indicated. **e** Intensity plot of *Lk* topoisomers visible in the 2D gel-blot. Colors and letters correspond to those in the 2D scheme. The x-axis indicates ΔLk relative to Lk^0 . The ΔLk of the minichromosome (mean \pm s.d., $n = 4$) is the difference between Lk^0 and Lk^{CHR} . See Supplementary Fig. 2 for detailed calculation of ΔLk . **f** The gel-blot (left) compares the *Lk* distribution of the YCp1.3 minichromosome extracted from fixed cells (lane 1) and that of the YCp1.3 minichromosome solubilized from non-fixed cells (lane 2). Intensity scans of lanes 1 and 2 are shown. The ethidium-stained gel (right) shows DNA relaxation activity in the solubilized chromatin. Supercoiled (S), relaxed (R) and nicked (N) forms of a reporter plasmid are indicated. **g** The gel-blot compares the *Lk* distribution of the YCp1.3 minichromosome in yeast cells cultured in rich medium (lane 1) and synthetic dropout medium (lane 2); and in yeast $\Delta top1$ (lane 3) and $\Delta top1 top2-4$ (lane 4) mutants. Intensity scans of lanes 1–4 are shown

topoisomerase at the same temperature (26 °C) used to generate the *Lk* of the YCp1.3 minichromosome in vivo (Fig. 1c). We analyzed the DNA samples using one- or two-dimensional (1D and 2D) agarose gel electrophoresis⁴⁷, in which we adjusted the concentrations of chloroquine in order to resolve in the same gel all the *Lk* topoisomers of YCp1.3 in vivo and that of its DNA relaxed in vitro. As seen in the 2D gel in Fig. 1d, the minichromosome (lane 1) and the relaxed DNA (lane 2) presented discrete distributions of *Lk* topoisomers (spots). Such *Lk* distributions occur because the energy difference between the *Lk* topoisomers is less than the thermal energy. The possible *Lk* topoisomers follow a Boltzmann distribution, whose means are Lk^0 for the relaxed DNA and Lk^{CHR} for the minichromosome DNA. We subtracted these values and found that YCp1.3 had a ΔLk of -5.81 (Fig. 1e, see Supplementary Fig. 2 for detailed calculation of ΔLk).

We next asked whether the ΔLk value of YCp1.3 was fully constrained by its chromatin structure or could be partially unconstrained (i.e., DNA supercoiling produced by gene transcription). To this end, we examined the topology of YCp1.3 when the minichromosome was solubilized from lysates of unfixed cells (Fig. 1f). In these conditions, cellular topoisomerases present in the cell lysates were able to relax supercoiled DNA plasmids completely (Fig. 1f, right). However, this DNA relaxation activity did not alter the *Lk* distribution of the minichromosome (Fig. 1f, left). We observed also that the *Lk* distribution of YCp1.3 was unchanged when yeast cells were cultured in rich medium and synthetic dropout medium, and

when YCp1.3 was hosted in yeast cells with reduced topoisomerase activity ($\Delta top1$ and $\Delta top1 top2-ts$) (Fig. 1g). All these observations indicated that the ΔLk value of -5.81 was fully constrained by the chromatin structure of YCp1.3. As yeast point centromeres restrain $+0.6$ units of ΔLk ⁴⁵, the five nucleosomes of YCp1.3 had to stabilize -6.4 units, an average ΔLk of -1.28 per nucleosome.

Insertion of a mononucleosome library into minichromosomes.

The average ΔLk of -1.28 per nucleosome in YCp1.3 assumes that all the minichromosomes are evenly occupied by nucleosomes I to V. However, native yeast nucleosomes are occasionally found to be partially unfolded and invading neighboring nucleosome territories or completely missing⁴⁸. Therefore, the absolute ΔLk per nucleosome could be larger than 1.28 if some nucleosomes were missing or unfolded. Conversely, this value could be smaller if the number of assembled nucleosomes were increased, although this possibility is less likely in light of the micrococcal nuclease data and the limited space available in YCp1.3 (Fig. 1a, b). However, we could not discard that chromatin elements other than nucleosomes could also contribute to the ΔLk value of YCp1.3. For instance, regulatory factors bound to the *TRP1* promoter and the *ARS1* region may alter the topology of the interacting DNA. Therefore, the average ΔLk value of -1.28 per nucleosome estimated above could not be accurate.

To obtain a more reliable ΔLk value, we conceived inserting an additional nucleosome into YCp1.3. The difference of ΔLk values

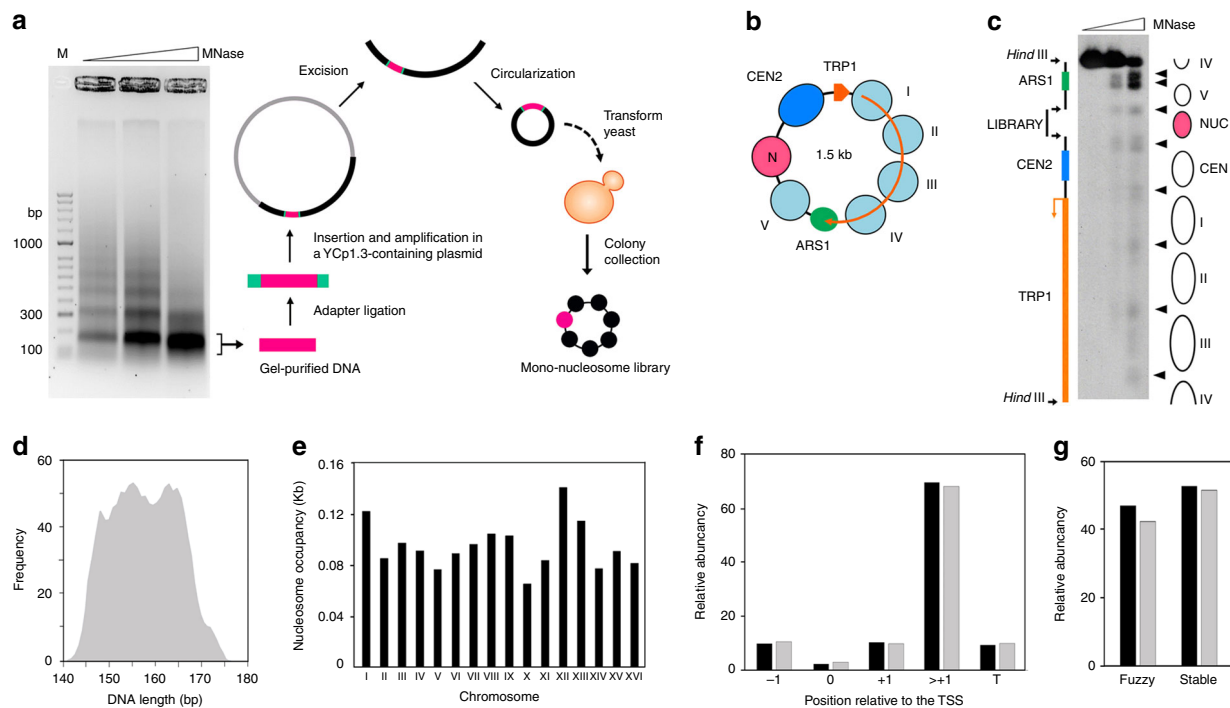


Fig. 2 Construction and structure of minichromosomes containing a mononucleosomal library. **a** Outline of the procedure used to construct circular minichromosomes containing a nucleosome library. **b** Scheme of the YCp1.3 minichromosome after the inclusion of the nucleosome library (red) between nucleosome V and CEN2. **c** Micrococcal nuclease digestion pattern of the bulk of YCp1.3 minichromosomes containing the library. Structural and functional elements of YCp1.3 and nuclease-sensitive sites (arrow heads) are indicated. **d** Length distribution of nucleosomal DNA fragments inserted in YCp1.3. **e** Chromosomal distribution of the nucleosomal library. The number of nucleosomal DNA sequences normalized per kb is plotted for each of the 16 chromosomes of *S. cerevisiae*. **f** Relative abundance of nucleosome types on the basis of their position relative to the transcription start site (TSS). Following the nomenclature in Jiang and Pugh (2009), “-1” is the first nucleosome upstream of the nucleosome-free region 5’ to TSS, “0” is a nucleosome that overlaps with the nucleosome-free region, “+1” is the first nucleosome downstream of the TSS, and “T” is the terminal nucleosome of a gene. **g** Relative abundance of nucleosomes on the basis of their positional stability. Nucleosomes are classified as “fuzzy or stable”, following Ioshikhes et al. (2006). Graphs in **f** and **g** compare relative abundance within our nucleosomal library (black) and within the genomic catalog of nucleosomes (gray)

(ΔLk) of YCp1.3 with and without this additional nucleosome would indicate the ΔLk stabilized by such a nucleosome. The average ΔLk constrained by individual nucleosomes in vivo could then be calculated by repeating this experiment with many different nucleosomes. To this end, we constructed a mono-nucleosome library as follows. We digested the whole chromatin of *S. cerevisiae* with increasing amounts of micrococcal nuclease to obtain ladders of nucleosomal DNA fragments. We purified the mononucleosomal DNA fragments (length ≈ 150 bp), added adapters, and inserted them into the YCp1.3 circle (Fig. 2a). In order not to interfere with the functional elements of YCp1.3, we allocated the nucleosome library insertions between nucleosome V and CEN2 (Fig. 2b, see Supplementary Fig. 3 for a detailed configuration of the insertion site). Upon transformation of YCp1.3 constructs carrying the mononucleosome library into yeast cells, we collected 1200 colonies. Micrococcal nuclease digestions of the minichromosomes pooled from all the colonies revealed a pattern of DNA cut sites that was nearly identical to that observed in native YCp1.3. However, a new protected DNA segment of about 150 bp appeared between nucleosome V and CEN2, consistent with the expected assembly of an additional nucleosome particle (Fig. 2c).

Parallel sequencing of the full library indicated that nearly all the colonies contained a distinct mononucleosomal DNA fragment inserted between nucleosome V and CEN2 of the YCp1.3 minichromosome. We mapped 1193 different sequences to the reference genome. Their average length was 156 ± 8 bp (mean \pm s.d.) (Fig. 2d). We identified them as previously

referenced nucleosomes by intersecting their coordinates with a catalog of nucleosome positions in yeast (Jiang and Pugh, 2009) (Supplementary Data 1). To determine whether our collection of nucleosomes was representative, we examined their chromosomal distribution (Fig. 2e), inter- or intra-genic position relative to transcription start sites⁶ (Fig. 2f), and positional stability⁴⁹ (Fig. 2g). The relative abundance of these nucleosome classes in our collection was comparable to that of the reference catalog. Therefore, the nucleosome library inserted in the YCp1.3 minichromosome was representative for the purpose of the intended analysis of nucleosomal DNA topology.

The average ΔLk value restrained by individual nucleosomes.

As with YCp1.3, we determined the ΔLk of minichromosomes carrying the nucleosome library by comparing their Lk distribution in vivo (Lk^{CHR}) with that of their relaxed DNA in vitro (Lk^0). Analysis of individual colonies of the library revealed that the minichromosomes had ΔLk values in the range of -7.0 to -7.1 (Fig. 3a). Therefore, relative to the ΔLk of -5.81 stabilized by YCp1.3, the inserted nucleosomes produced $\Delta \Delta Lk$ of -1.2 to -1.3 units. These values were consistent with the average ΔLk of -1.28 per nucleosome calculated for YCp1.3 (Fig. 1e). Moreover, the five clones analyzed in Fig. 3a represented nucleosomes of distinct allocation relative to TSS (-1 , $+1$, $>+1$) and different positional stability (fuzzy or stable). Therefore, these nucleosomes stabilized comparable ΔLk values irrespective of the nucleosome category. The above analysis of individual colonies also showed

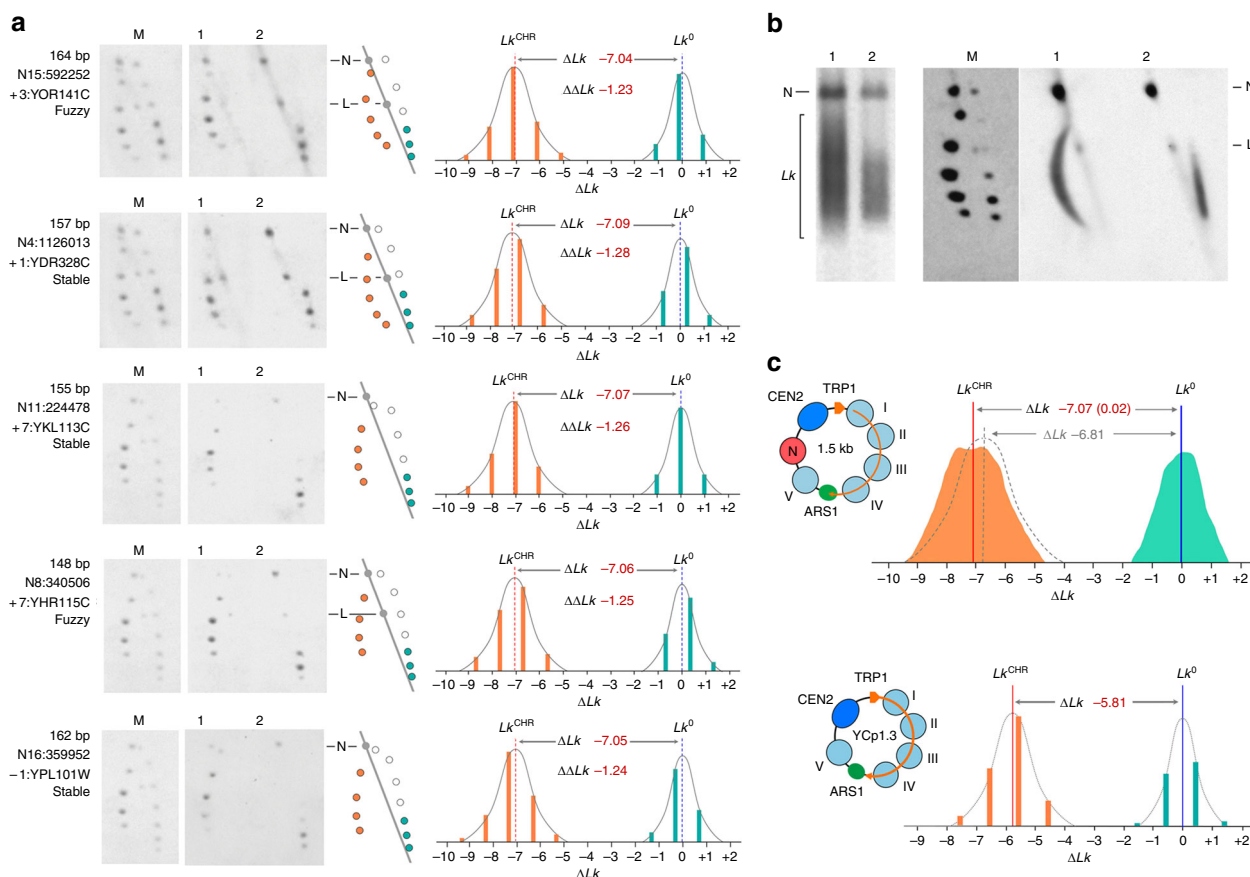


Fig. 3 ΔLk gain produced by the mononucleosome library. **a** DNA topology of five individual minichromosomes of the nucleosomal library. In each case, the length (bp) of the inserted mononucleosomal DNA fragment, the nucleosome ID, genic positioning and stability are indicated. 2D gel-blots show a marker of Lk topoisomers (lane M), the minichromosome DNA extracted from fixed cells (lane 1) and after its relaxation with topo I (lane 2). 2D schemes show the relative position of Lk topoisomers visible in lanes 1 (orange dots) and 2 (green dots). The intensity plots of the above topoisomers showing the mean of the Lk distributions (Lk^0 and Lk^{CHR}) and the resulting ΔLk value were obtained as in Fig. 1e. $\Delta\Delta Lk$ is the ΔLk gain produced relative to YCp1.3 ($\Delta Lk = -5.81$). **b** 1D and 2D gel-blots show the DNA of the pool of minichromosomes extracted from the entire library (lane 1) and the pooled DNA after its relaxation with topoisomerase I (lane 2). The 2D gel includes a marker of individual Lk topoisomers (lane M). Nicked (N) and linear (L) molecules are indicated. **c** Top, Lk distributions corresponding to the minichromosomes carrying the nucleosome library (orange) and their DNAs after relaxation with topo I (green). The mean of the pooled Lk distributions (Lk^0 and Lk^{CHR}) and the resulting ΔLk value of -7.07 ± 0.02 (mean \pm s.d.) from four replicate experiments were determined as detailed in Supplementary Fig. 4. The position of a hypothetical Lk distribution with ΔLk of -6.81 , which would have implied that the nucleosome library produced a $\Delta\Delta Lk$ of -1.0 , is illustrated (dashed gray). Bottom, Lk distributions and ΔLk of YCp1.3 are shown at the same scale and denote that $\Delta\Delta Lk = -1.26$

that although different minichromosomes had nearly the same ΔLk , the positions of their Lk topoisomers in the gel did not have the same phasing with respect to Lk^0 (Fig. 3a). This-misalignment of the Lk topoisomers occurred because the nucleosomal DNA fragments inserted had different lengths (mean 156 bp, s.d. ± 8 bp), and Lk phasing occurs only when the length differences are multiples of the helical repeat of DNA ($h \approx 10.5$ bp)²⁴. However, the length differences producing Lk misalignment were small compared to the size of the minichromosomes (about 1.57 Kb). Therefore, we were able to analyze the bulk of Lk distributions of the minichromosomes carrying the nucleosome library in a single electrophoresis run, rather than analyzing them individually. To this end, we pooled the colonies of the library to obtain all the Lk distributions of the minichromosomes in one DNA sample and all the Lk distributions of relaxed DNA in another.

As expected, 1D and 2D gel electrophoresis of the pooled samples did not reveal single bands of Lk topoisomers but smeary signals as a result of the overlapping of numerous Lk distributions (Fig. 3b). These overlapped signals presented small protrusions

(Fig. 3c, green and orange), which could suggest that the pooled Lk distributions were bimodal. This outcome would occur if near one half the nucleosomes of the library constrained $\Delta Lk \approx -1.0$ and the other half $\Delta Lk \approx -1.5$. However, this scenario is not consistent with the data of individual nucleosomes, all which constrained ΔLk values between -1.2 and -1.3 (Fig. 3a). Actually, these protrusions were expected for another reason. They appeared because the different lengths of the DNA inserts were not equally represented (Fig. 2d) and thus so was the phasing of their corresponding LK topoisomers. Accordingly, some protrusions appeared also in the pool of relaxed DNAs. The fact that the pooled Lk distributions presented a dispersion similar to that of individual Lk distributions further substantiated that the pooled samples were essentially monomodal. In the case of the relaxed DNA (Fig. 3c, green), the similar dispersion of the pooled and individual Lk distributions corroborated that the gel position of Lk^0 was virtually the same for all the DNA molecules of the library regardless of small differences in length. In the case of the minichromosomes (Fig. 3c, orange), the analogous dispersion of the pooled and individual Lk

distributions implied that most minichromosomes had nearly the same Lk^{CHR} value. Consequently, most nucleosomes of the library produced the same $\Delta\Delta Lk$ with respect to the ΔLk of the YCp1.3 minichromosome. From the means (Lk^{CHR} and Lk^0 of the pooled Lk distributions), we found that the minichromosomes containing the nucleosome library had a ΔLk of -7.07 (see Supplementary Fig. 4 for detailed calculation). This ΔLk value, which represented a $\Delta\Delta Lk$ of -1.26 relative to the ΔLk of YCp1.3, was consistent in four replicate experiments (s.d. ± 0.02) and in agreement with that of individual minichromosomes (Fig. 3a). Therefore, we concluded that -1.26 is the average value of the ΔLk stabilized by individual nucleosomes. Note that, if the inserted library of nucleosomes had restrained $\Delta\Delta Lk$ values of about -1.0 , the minichromosomes would have presented an average ΔLk of -6.81 , implying a noticeable displacement of the Lk^{CHR} position in the intensity plot (Fig. 3c).

Discussion

Here we show experimental evidence that provides a solution to the long-standing Lk paradox of nucleosomal DNA. Our results indicate that most native nucleosomes constrain a ΔLk close to -1.26 . This value differs markedly from the generally assumed ΔLk value of -1.0 , which was established in earlier studies. We believe that this discrepancy is due to the distinct chromatin models and limited accuracy of the procedures that were used previously to estimate the ΔLk constrained by nucleosome particles.

One source of inaccuracy was in determining the exact number of nucleosomes assembled in circular DNA molecules. Previous studies using SV40 as a chromatin model relied on electron microscopy for counting nucleosomes or nucleosome-like particles. The numbers obtained by different laboratories varied from 20 to 27^{17–19,50}. Recent mapping of nucleosome positions in the SV40 genome has revealed that this variability is not only experimental. Intracellular SV40 minichromosomes and SV40 virions present variable nucleosome number and epigenetic modifications that alter the nucleosome organization depending on the infection stage⁵¹. This variability in nucleosome number could therefore have produced imprecise ΔLk values, especially when nucleosome counting and DNA topology analyses were done with uncorrelated samples and by different laboratories^{15,16}. Not surprisingly, the SV40 model supported a broad range of ΔLk values, including -1.25 per nucleosome¹⁶. The uncertainty in the exact number of nucleosomes present in circular DNA molecules also affected studies using chromatin reconstitution in vitro, which also relied on electron microscopy for counting nucleosomes^{21,22}. Moreover, in these studies, chromatin reconstitution in tandem repeats of nucleosome positioning sequences could have markedly deviated the ΔLk values per nucleosome (from -0.8 to -1.4) depending on the periodic length assigned to DNA linker segments³⁸.

The other source of imprecision in earlier studies was in the calculation of ΔLk from the DNA bands observed in agarose gels. In most studies using SV40 and reconstituted chromatin, the gel position of Lk^0 was often approximated to that of the slowest Lk topoisomer^{15,19,21,22,50}, instead of being allocated to the mean Lk of the relaxed Lk distribution¹⁶. Likewise, in earlier measurements using circular minichromosomes of yeast, the gel position of Lk^0 was taken as that of nicked DNA circles. This was the case of the *TRPIARS1* minichromosome (1.45 kb), which contains seven nucleosomes and was assigned a ΔLk of -720 . Finally, in most previous studies, it was unclear whether the processing of chromatin samples (to determine Lk^{CHR}) and the relaxation of naked DNA circles (to determine Lk^0) were quenched at the same temperature. Since the helical repeat of

DNA lessens as the temperature diminishes⁴⁶, quenching the topology of DNA at 4°C produces Lk values up to 1 unit/kb higher than at 37°C ⁴⁵.

Our experimental approach minimized the uncertainty in nucleosome counting and ΔLk calculation. The small minichromosomes used presented well-defined nucleosome positions, which were bounded by specific chromatin elements (*TRP1* promoter, *ARS1*, *CEN2*). The small size also circumvented significant effects of high order folding of the chromatin on the ΔLk values. Our experimental results corroborated the ΔLk value stabilized by individual nucleosomes in two ways. First, by averaging the ΔLk of the minichromosomes by their number of nucleosomes, we obtained a ΔLk of -1.28 per nucleosome. However, this measurement did not take into account plausible variability in nucleosome occupancy and effects of structural elements other than nucleosomes. We reduced these ambiguities by determining the Lk gain ($\Delta\Delta Lk$) produced upon the insertion of the nucleosome library. We obtained thereby the more reliable ΔLk value of -1.26 per nucleosome. We found also that there is very little dispersion in the ΔLk constrained by the nucleosome library, which indicated that the majority of nucleosomes stabilize a similar DNA topology. Our results could be hardly explained if native nucleosomes were each stabilizing a ΔLk of -1.0 . This value could stand if the minichromosomes had assembled a number of nucleosomes higher than expected, which seems unlikely in light of the micrococcal nuclease data and the space available. A ΔLk of -1.0 per nucleosome could also stand if the minichromosomes were spatially compacted by adopting a strong negative writhe (i.e., $\Delta Wr \approx -1.0$). Such folding would imply that DNA linker lengths and the subsequent rotational orientations between adjacent nucleosomes are alike in all minichromosomes. However, the inserted nucleosome library comprised segments of various lengths and the resulting minichromosomes still constrained very similar ΔLk values.

The ΔLk value of -1.26 leads to a reevaluation of the Lk paradox of nucleosomal DNA in terms of the general equation $\Delta Lk = \Delta Tw + \Delta Wr$ ¹⁴. Considering that the core DNA is globally overtwisted by about $+0.2$ turns ($\Delta Tw \approx +0.2$)^{12,35}, the stabilization of -1.26 units of ΔLk implies that the writhe acquired (ΔWr) by DNA upon nucleosome formation should be about -1.46 . The Wr of DNA in mononucleosomes has been computed to be around -1.5 ^{11,36}. Here we calculated this value for different degrees of superhelical turning of core DNA by $Wr = n(1 - \sin\theta)$ ^{14,52}, where n is the number of helical turns and θ is the pitch angle of the turns (supplementary Fig 5). Nucleosomal Wr is about -1.53 when the core DNA completes 1.65 left-handed superhelical turns around the histone octamer. This conformation corresponds to that of the crystallized nucleosome structure¹ and also to that of chromatosomes⁵³, in which the entry and exit DNA linker segments cross in an angle of about 60° that is fixed by histone H1. This Wr value is likely to reflect thus the upper limit of the absolute DNA writhe of nucleosomes in solution. However, yeast has very low level of linker histone⁵⁴, though nuclease digestions support the existence of proto-chromatosome structures⁵⁵. Moreover, extensive experimental evidence has demonstrated that the conformational dynamics of nucleosomes in physiological conditions frequently leads to partial unwrapping or breathing motions of the core DNA^{56–59}. These motions substantially reduce the absolute Wr of the nucleosomal DNA and thereby its average value. For instance, just by reducing the wrapping of DNA to 1.5 left-handed superhelical turns, mononucleosomal Wr drops to -1.38 (supplementary Fig 5). Therefore, an average ΔWr of about -1.46 per nucleosome is a realistic topological mark, which along the ΔTw of about $+0.2$ and the ΔLk of -1.26 , provides a plausible explanation for the linking number paradox of nucleosomal DNA.

Our experimental findings contribute to a better understanding of how DNA supercoiling energy is confined by nucleosomes, and of how nucleosomes buffer the DNA supercoiling generated during gene transcription. Our experimental approach also leads to a new genome-wide categorization of nucleosomes on the basis of their DNA topology, thus opening a new dimension toward deciphering the mechanisms that orchestrate chromatin structure and functions.

Methods

Construction of minichromosomes and the nucleosome library. To construct YCp1.3 (1341 bp), a 997 bp genomic segment of *S. cerevisiae* containing *TRP1-ARSI* (coordinates 461739 to 462736) and a 243 bp genomic segment containing *CEN2* (coordinates 238194 to 238437) were amplified by PCR. Both segments were ligated and inserted into a plasmid vector via endonuclease restriction sites engineered by PCR. Subsequent digestion with endonuclease *NotI* released a 1341-bp fragment containing the *TRP1-ARSI-CEN2* sequence. This fragment was circularized with T4 DNA ligase and monomeric circles were gel-purified to obtain the YCp1.3 circle. See Figure S1 for a description of the oligonucleotides used for PCR and the complete bp sequence of YCp1.3. The YCp1.3 circle was used to transform the *S. cerevisiae* strain FY251 (*MATa his3-Δ200 leu2-D1 trp1-Δ63 ura3-52*) and its topoisomerase-mutant derivatives JCW27 (*Δtop1*) and JCW28 (*Δtop1 top2-4*)⁶⁰. To construct the nucleosome library, yeast cells from a 250 ml culture (OD 1.0) were collected, washed with water, and incubated with 80 ml of 1 M Sorbitol, 30 mM DTT for 15 min at 28 °C. Next, 625 U of Lyticase (Sigma-Aldrich L2524) and 10 μL of 4 M NaOH were added to the cells suspension, and the incubation continued until > 80% of cells were converted into spheroplasts. Spheroplasts were washed with 1M Sorbitol and resuspended in 1.5 ml of hypotonic lysis buffer (1 mM CaCl₂, 5 mM KH₂PO₄, 1 mM PMSF) at 24 °C. Next, 30 units of micrococcal nuclease (Sigma-Aldrich N3755) were added, and the mixture was incubated at 24 °C. Aliquots of 300 μl were quenched with 20 mM EDTA 1% SDS at different time points (3–30 min). The digestion of chromosomal DNA was examined by gel electrophoresis (1% agarose in TBE buffer, 80 V during 3 h). Mononucleosomal DNA fragments (about 150 bp in length) produced at different time points were gel-eluted and pooled. The severed DNA fragments produced by micrococcal nuclease were repaired by removing terminal 3'-phosphates with T4-polynucleotide kinase and by filling the DNA ends with Klenow and T4-DNA polymerase activities. The nucleosomal DNA fragments were A-tailed with Klenow activity and ligated to adapters. The adapters included an *AscI* site, which permitted the insertion of the mononucleosomal DNA fragments into the single *AscI* site of YCp1.3. See Figure S2 for a description of the adapters and the site of insertion in YCp1.3. The YCp1.3 constructs containing the library of mononucleosomal DNA sequences were amplified in bacterial plasmids. Upon digestion with endonuclease *NotI*, the fragments of about 1.57 Kb containing the library within YCp1.3 were circularized with T4 DNA ligase. Monomeric circles were gel-purified and used to transform FY251.

Yeast culture and DNA extraction of fixed minichromosomes. Yeast cells transformed with YCp1.3 were grown at 26 °C in standard yeast synthetic media containing TRP dropout supplement (Sigma) with 2% glucose or in rich YPD medium, as indicated. When the liquid cultures (20 ml) reached mid-log phase (OD ≈ 0.8), the cells were fixed *in vivo* by quickly mixing the cultures with one cold volume (–20 °C) of ET solution (Ethanol 95%, Toluene 28 mM, Tris-HCl pH 8.8 20 mM, EDTA 5 mM). The following steps were done at room temperature. Cells were sedimented, washed twice with water, resuspended in 400 μl of TE, and transferred to a 1.5-ml microfuge tube containing 400 μl of phenol and 400 μl of acid-washed glass beads (425–600 μm, Sigma). Mechanic lysis of >80% cells was achieved by shaking the tubes in a FastPrep® apparatus for 10 s at power 5. The aqueous phase of the lysed cell suspension was collected, extracted with chloroform, precipitated with ethanol, and resuspended in 100 μl of TE containing RNase-A. After a 15-min incubation at 37 °C, the samples were extracted with phenol and chloroform, DNA precipitated with ethanol and resuspended in 30 μl of TE. The same procedure of cell culture, *in vivo* fixation, cell lysis and DNA extraction was applied to individual colonies of the minichromosome library. In the case of sampling the full library, the bulk of colonies were collected from agar plates, washed with water, and diluted (OD 0.2) in 200 ml of standard yeast synthetic media (TRP dropout). The pooled cells were grown at 26 °C and fixed when the liquid cultures reached mid-log phase.

Yeast culture and solubilization of minichromosomes. Liquid yeast cultures (20 ml) at mid-log phase were sedimented, washed with water, and resuspended in 500 μl of buffer L (Tris-HCl 10 mM pH 8, EDTA 1 mM, EGTA 1 mM, DTT 1 mM, NaCl 150 mM, Triton 0.1%, pepstatin 1 mg/ml, leupeptin 1 mg/ml, and PMSF 1 mM). The suspension was transferred to a 1.5-ml microfuge tube containing 500 μl of acid-washed glass beads (425–600 μm, Sigma). Mechanical lysis of >80% cells was achieved after six cycles of 30 s of vortexing plus 30 s of ice cooling. The

supernatant of the lysate was recovered by centrifugation (2000 × g) and loaded on a Sephacryl S-300 column equilibrated with buffer L at 4 °C. The first filtration volume containing the circular minichromosomes was recovered and incubated at 26 °C for 10 min. A supercoiled DNA plasmid was added to one aliquot of the eluted minichromosomes and further incubated for 5 min in order to test the DNA relaxation activity of cellular topoisomerases. The reactions were then quenched with one volume of 20 mM EDTA, 1% SDS, and 100 mg/ml proteinase K, followed by an incubation at 60 °C for 30 min. The samples were extracted with phenol and chloroform, and the DNA was precipitated with ethanol and resuspended in 30 μl of TE.

Micrococcal nuclease mapping of chromatin structure.

YCp1.3 minichromosomes and their derivatives were solubilized and eluted from a Sephacryl S-300 column as described above. Eluted minichromosomes were adjusted to 2 mM CaCl₂ and pre-incubated at 25 °C for 5 min. Micrococcal nuclease was added (2–100 units/ml), and incubations proceeded at 25 °C for 5 min. The digestions were quenched with one volume of 20 mM EDTA, 0.5% SDS, and 100 mg/ml proteinase K, followed by incubation at 60 °C for 30 min. The mixtures were extracted with phenol and chloroform, and the DNA precipitated with ethanol and resuspended in 30 μl of TE. The DNA was then digested with *HindIII* restriction endonuclease, which has a single cutting site in YCp1.3. The resulting DNA fragments were separated by gel electrophoresis (1.2% agarose), blotted, and probed with a short DNA sequence (194 bp) starting at the *HindIII* site of YCp1.3.

DNA sequencing and analysis. DNA extracted from minichromosomes containing the mononucleosome library was sequenced (Illumina HiSeq 2000, 50 base paired-end reads), and resulting FASTQ data files were subject to QC using Cutadapt (1.12). Sequences were then mapped to the *S. cerevisiae* reference genome (SacCer3) using bowtie (v1.1.2). Once nucleosome coordinates had been established, subsequent analyses were performed by integrating published data sets (Ioshikhes et al 2006; Jiang and Pugh, 2009) and by using bedtools (v2.27) and Galaxy.

DNA relaxation with topoisomerase I. DNA purified from minichromosome preparations was pre-incubated at 26 °C for 5 min in 30 μl of Tris-HCl 10 mM pH 8, EDTA 1 mM, and NaCl 150 mM. Catalytic amounts of topoisomerase I of vaccinia virus⁴⁵ were then added, and the incubations proceeded at 26 °C for 30 min. Reactions were quenched with one volume of 20 mM EDTA and 1% SDS.

Electrophoresis of Lk topoisomers. DNA from YCp1.3 (1341 bp) and from minichromosomes containing the mononucleosome library (about 1.57 kb) were loaded onto 1.4% (w/v) agarose gels. One-dimensional electrophoresis was carried out at 2.5 V/cm for 18 h in TBE buffer (89 mM Tris-borate and 2 mM EDTA) containing 0.55 μg/ml chloroquine. Two-dimensional electrophoresis was in TBE containing 0.55 μg/ml chloroquine in the first dimension (2.5 V/cm for 18 h) and in TBE containing 3 μg/ml chloroquine in the second dimension (5 V/cm for 4 h). Gels were blot-transferred to a nylon membrane and probed at 60 °C with the *TRPIARS1* sequence labeled with *AlkPhos* Direct (GE Healthcare®). Chemiluminescent signals of increasing exposure periods were recorded on X-ray films and non-saturated signals of individual *Lk* topoisomers and bins of pooled *Lk* distributions quantified with ImageJ.

Lk distribution analysis and calculation of ΔLk. In the case of individual minichromosomes, the most intense topoisomer of the *Lk* distribution of relaxed DNA was initially assigned the value $\Delta Lk = 0$. An Integer ΔLk value (positive or negative) was subsequently assigned to each topoisomer of the *Lk* distributions of minichromosome and relaxed DNA according to the *Lk* markers included in the 2D gels. The mean value of each *Lk* distribution (Lk^0 and Lk^{CHR}) was calculated, and the ΔLk between the *Lk* distributions was obtained by subtracting their means (see details in Supplementary Fig. 2). In the case of pooled minichromosomes, continuous *Lk* distributions were quantified by bins and their mean calculated. The gel position of the means was interpolated with the that of *Lk* marker to obtain ΔLk values of the means in the marker frame. The ΔLk between pooled minichromosomes and relaxed DNAs was determined by subtracting their means (see details in Supplementary Fig. 4). In all figures, the *Lk* results were illustrated by plotting the intensity of *Lk* topoisomers of minichromosomes and relaxed DNA along a scale of ΔLk units (*x*-axis), in which the value $\Delta Lk = 0$ was re-adjusted to the mean of the *Lk* distribution of the relaxed DNA (Lk^0).

Data availability

The data that support the findings of this study are available from the corresponding author upon request.

Received: 13 March 2018 Accepted: 11 September 2018

Published online: 28 September 2018

References

- Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**, 251–260 (1997).
- Zlatanova, J., Bishop, T. C., Victor, J. M., Jackson, V. & van Holde, K. The nucleosome family: dynamic and growing. *Structure* **17**, 160–171 (2009).
- Rando, O. J. & Chang, H. Y. Genome-wide views of chromatin structure. *Annu. Rev. Biochem.* **78**, 245–271 (2009).
- Andrews, A. J. & Luger, K. Nucleosome structure(s) and stability: variations on a theme. *Annu. Rev. Biophys.* **40**, 99–117 (2011).
- Bai, L. & Morozov, A. V. Gene regulation by nucleosome positioning. *Trends Genet.* **26**, 476–483 (2010).
- Jiang, C. & Pugh, B. F. Nucleosome positioning and gene regulation: advances through genomics. *Nat. Rev. Genet.* **10**, 161–172 (2009).
- Radman-Livaja, M. & Rando, O. J. Nucleosome positioning: how is it established, and why does it matter? *Dev. Biol.* **339**, 258–266 (2010).
- Crick, F. H. Linking numbers and nucleosomes. *Proc. Natl Acad. Sci. USA* **73**, 2639–2643 (1976).
- Finch, J. T. et al. Structure of nucleosome core particles of chromatin. *Nature* **269**, 29–36 (1977).
- Wang, J. C. The path of DNA in the nucleosome. *Cell* **29**, 724–726 (1982).
- Prunell, A. A topological approach to nucleosome structure and dynamics: the linking number paradox and other issues. *Biophys. J.* **74**, 2531–2544 (1998).
- Bates, A. D. & Maxwell, A. (eds) in *DNA Topology* 83–103 (Oxford University Press, 2005).
- Marko, J. F. & Siggia, E. D. Statistical mechanics of supercoiled DNA. *Phys. Rev. E. Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.* **52**, 2912–2938 (1995).
- Fuller, F. B. The writhing number of a space curve. *Proc. Natl Acad. Sci. USA* **68**, 815–819 (1971).
- Keller, W. Determination of the number of superhelical turns in simian virus 40 DNA by gel electrophoresis. *Proc. Natl Acad. Sci. USA* **72**, 4876–4880 (1975).
- Shure, M. & Vinograd, J. The number of superhelical turns in native virion SV40 DNA and minicoll DNA determined by the band counting method. *Cell* **8**, 215–226 (1976).
- Saragosti, S., Moyne, G. & Yaniv, M. Absence of nucleosomes in a fraction of SV40 chromatin between the origin of replication and the region coding for the late leader RNA. *Cell* **20**, 65–73 (1980).
- Sogo, J. M., Stahl, H., Koller, T. & Knippers, R. Structure of replicating simian virus 40 minichromosomes. The replication fork, core histone segregation and terminal structures. *J. Mol. Biol.* **189**, 189–204 (1986).
- Germond, J. E., Hirt, B., Oudet, P., Gross-Bellard, M. & Chambon, P. Folding of the DNA double helix in chromatin-like structures from simian virus 40. *Proc. Natl Acad. Sci. USA* **72**, 1843–1847 (1975).
- Pederson, D. S., Venkatesan, M., Thoma, F. & Simpson, R. T. Isolation of an episomal yeast gene and replication origin as chromatin. *Proc. Natl Acad. Sci. USA* **83**, 7206–7210 (1986).
- Simpson, R. T., Thoma, F. & Brubaker, J. M. Chromatin reconstituted from tandemly repeated cloned DNA fragments and core histones: a model system for study of higher order structure. *Cell* **42**, 799–808 (1985).
- Norton, V. G., Imai, B. S., Yau, P. & Bradbury, E. M. Histone acetylation reduces nucleosome core particle linking number change. *Cell* **57**, 449–457 (1989).
- Klug, A. & Lutter, L. C. The helical periodicity of DNA on the nucleosome. *Nucleic Acids Res.* **9**, 4267–4283 (1981).
- Wang, J. C. Helical repeat of DNA in solution. *Proc. Natl Acad. Sci. USA* **76**, 200–203 (1979).
- White, J. H., Cozzarelli, N. R. & Bauer, W. R. Helical repeat and linking number of surface-wrapped DNA. *Science* **241**, 323–327 (1988).
- Noll, M. Internal structure of the chromatin subunit. *Nucleic Acids Res.* **1**, 1573–1578 (1974).
- Lutter, L. C. Precise location of DNase I cutting sites in the nucleosome core determined by high resolution gel electrophoresis. *Nucleic Acids Res.* **6**, 41–56 (1979).
- Prunell, A. et al. Periodicity of deoxyribonuclease I digestion of chromatin. *Science* **204**, 855–858 (1979).
- Hayes, J. J., Tullius, T. D. & Wolffe, A. P. The structure of DNA in a nucleosome. *Proc. Natl Acad. Sci. USA* **87**, 7405–7409 (1990).
- Hayes, J. J., Clark, D. J. & Wolffe, A. P. Histone contributions to the structure of DNA in the nucleosome. *Proc. Natl Acad. Sci. USA* **88**, 6829–6833 (1991).
- Drew, H. R. & Travers, A. A. DNA structural variations in the *E. coli* tyrT promoter. *Cell* **37**, 491–502 (1984).
- Satchwell, S. C., Drew, H. R. & Travers, A. A. Sequence periodicities in chicken nucleosome core DNA. *J. Mol. Biol.* **191**, 659–675 (1986).
- Widom, J. Short-range order in two eukaryotic genomes: relation to chromosome structure. *J. Mol. Biol.* **259**, 579–588 (1996).
- White, J. H. & Bauer, W. R. Calculation of the twist and the writhe for representative models of DNA. *J. Mol. Biol.* **189**, 329–341 (1986).
- White, J. H. & Bauer, W. R. The helical repeat of nucleosome-wrapped DNA. *Cell* **56**, 9–10 (1989).
- Le Bret, M. Computation of the helical twist of nucleosomal DNA. *J. Mol. Biol.* **200**, 285–290 (1988).
- Worcel, A., Strogatz, S. & Riley, D. Structure of chromatin and the linking number of DNA. *Proc. Natl Acad. Sci. USA* **78**, 1461–1465 (1981).
- Nikitina, T., Norouzi, D., Grigoryev, S. A. & Zhurkin, V. B. DNA topology in chromatin is defined by nucleosome spacing. *Sci. Adv.* **3**, e1700957 (2017).
- De Lucia, F., Alilat, M., Sivolob, A. & Prunell, A. Nucleosome dynamics. III. Histone tail-dependent fluctuation of nucleosomes between open and closed DNA conformations. Implications for chromatin dynamics and the linking number paradox. A relaxation study of mononucleosomes on DNA minicircles. *J. Mol. Biol.* **285**, 1101–1119 (1999).
- Bancaud, A. et al. Structural plasticity of single chromatin fibers revealed by torsional manipulation. *Nat. Struct. Mol. Biol.* **13**, 444–450 (2006).
- White, C. L., Suto, R. K. & Luger, K. Structure of the yeast nucleosome core particle reveals fundamental changes in internucleosome interactions. *EMBO J.* **20**, 5207–5218 (2001).
- Rando, O. J. & Winston, F. Chromatin and transcription in yeast. *Genetics* **190**, 351–387 (2012).
- Thoma, F., Bergman, L. W. & Simpson, R. T. Nuclease digestion of circular TRP1ARS1 chromatin reveals positioned nucleosomes separated by nuclease-sensitive regions. *J. Mol. Biol.* **177**, 715–733 (1984).
- Jiang, C. & Pugh, B. F. A compiled and systematic reference map of nucleosome positions across the *Saccharomyces cerevisiae* genome. *Genome Biol.* **10**, R109 (2009).
- Diaz-Ingelmo, O., Martinez-Garcia, B., Segura, J., Valdes, A. & Roca, J. DNA topology and global architecture of point centromeres. *Cell Rep.* **13**, 667–677 (2015).
- Depew, D. E. & Wang, J. C. Conformational fluctuations of DNA helix. *Proc. Natl Acad. Sci. USA* **72**, 4275–4279 (1975).
- Hanai, R. & Roca, J. Two-dimensional agarose-gel electrophoresis of DNA topoisomers. *Methods Mol. Biol.* **94**, 19–27 (1999).
- Chereji, R. V. & Morozov, A. V. Ubiquitous nucleosome crowding in the yeast genome. *Proc. Natl Acad. Sci. USA* **111**, 5236–5241 (2014).
- Ioshikhes, I. P., Albert, I., Zanton, S. J. & Pugh, B. F. Nucleosome positions predicted through comparative genomics. *Nat. Genet.* **38**, 1210–1215 (2006).
- Germond, J. E. Simian virus 40 DNA extracted from infected cells with sodium deoxycholate no longer reflects its in vivo superhelix density. *J. Virol.* **26**, 681–685 (1978).
- Kumar, M. A. et al. Nucleosome positioning in the regulatory region of SV40 chromatin correlates with the activation and repression of early and late transcription during infection. *Virology* **503**, 62–69 (2017).
- Vologodskii, A. V. & Cozzarelli, N. R. Conformational and thermodynamic properties of supercoiled DNA. *Annu. Rev. Biophys. Biomol. Struct.* **23**, 609–643 (1994).
- Bednar, J. et al. Structure and dynamics of a 197 bp nucleosome in complex with linker histone H1. *Mol. Cell* **66**, 384–397 (2017). e8.
- Panday, A. & Grove, A. Yeast HMO1: linker histone reinvented. *Microbiol. Mol. Biol. Rev.* **81**, 1–21 (2017).
- Cole, H. A. et al. Novel nucleosomal particles containing core histones and linker DNA but no histone H1. *Nucleic Acids Res.* **44**, 573–581 (2016).
- Li, G. & Widom, J. Nucleosomes facilitate their own invasion. *Nat. Struct. Mol. Biol.* **11**, 763–769 (2004).
- Tomschik, M., Zheng, H., van Holde, K., Zlatanova, J. & Leuba, S. H. Fast, long-range, reversible conformational fluctuations in nucleosomes revealed by single-pair fluorescence resonance energy transfer. *Proc. Natl Acad. Sci. USA* **102**, 3278–3283 (2005).
- Li, G., Levitus, M., Bustamante, C. & Widom, J. Rapid spontaneous accessibility of nucleosomal DNA. *Nat. Struct. Mol. Biol.* **12**, 46–53 (2005).
- Bucceri, A., Kapitzka, K. & Thoma, F. Rapid accessibility of nucleosomal DNA in yeast on a second time scale. *EMBO J.* **25**, 3123–3132 (2006).
- Trigueros, S. & Roca, J. Failure to relax negative supercoiling of DNA is a primary cause of mitotic hyper-recombination in topoisomerase-deficient yeast cells. *J. Biol. Chem.* **277**, 37207–37211 (2002).

Acknowledgements

This work was supported by grants from Plan Estatal de Investigación Científica y Técnica of Spain to J.R. (BFU2015-67007-P and MDM-2014-0435-02). J.R. is a Professor at the Spanish National Research Council (CSIC).

Author contributions

J.R. conceived the project. J.R. and J.S. designed experiments. J.S., O.D., B.M., A.V. and S.D. performed experiments. J.R., J.S. and R.S.J. analyzed data. J.R. wrote the manuscript.

Additional information

Supplementary Information accompanies this paper at <https://doi.org/10.1038/s41467-018-06547-w>.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in

published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018

Supplementary Information

Intracellular nucleosomes constrain a DNA linking number difference of -1.26 that reconciles the *Lk* paradox

Segura et al.

A

Oligonucleotides to amplify genomic coordinates 461739 to 462736 (997 bp) that include TRP-ARS1

F: 5' GATAGAATTTCGGTCGAAAAAAGAAAAGGAGAGGGCCAAGAGGGAG 3'
R: 5' AAGGCGCGCCAGCGGCCATTCTCTCAACAAGTTTGATTCC 3'

Oligonucleotides to amplify genomic coordinates 238194 to 238437 (243 bp) that include CEN2

F: 5' TATAGGATCCCATACATGAAGAATATAGGATCTGTATATCG 3'
R: 5' CATAGGTACCTAATAATAAATTAATCTTGAGCAAATTGGTCC 3'

B

YCp1.3 (1341 bp)

```
CGGGCCGC TGGCGGCC GGATCC CATACATGAAGAATATAGGATCTGTATATCGTCGTCACCGACGTAAGGTTTTTTTGT
CAATTC TCGGTGAACAATTTTTGCAAGAAATATATTGATACCTCTTGTGAAATATTTTTCTTT TTCGGAAACAAAAA
CATAAAAATAATTATTTAAGAAAAAAATTAATTAATAAATTTAAATAGAAAAATCTTTTTTACTTAATAATAATTAATAA
ATAAGT CACATGAT TTATGTAGGACCAATTTGCTCAAGATTAATTTATTATTA GGTACCAAGATAATAGTAACCCGGGT
AATAAGTGTATTTGAGATA GAATTCGGTCGAAAAAAGAAAAGGAGAGGGCCAAGAGGGAGGGCATTGGTGACTATTGAGC
ACGTGAGTATACGTGATTAAGCACACAAGGCAGCTTGGAGTATGCTCTGTTATTAATTTACAGGTAGTTCTGGTCCATT
GGTGAAAGTTGCGGCTTGACAGACACAGAGGCCGAGAAATGTGCTCTAGATTCCGATGCTGACTTGCTGGGTATTATATG
TGTGCCCAATAGAAAGAGAACAATTGACCCGGTTATTGCAAGGAAAAATTTCAAAGTCTTGTAAGCATATAAAAAATAGT
TCAGGCACCTCCGAAATACTTGGTTGGCGTGTTCGTAATCAACCTAAGGAGGATGTTTGGCTCTGGTCAATGATTACGG
CATTGATATCGTCCAAC TGCATGGAGATGAGTCGTGGCAAGAAATACCAAGAGTTCCTCGGTTTGCCAGTTATTAAGAGAC
TCGTATTTCCAAAAGACTGCAACATACTACTCAGTGCAGCTTCACAGAAACCTCATTTCGTTTATCCCTTGTTTGATTCA
GAAGCAGGTGGGACAGGTGAAC TTTGGATTGGAAC TCGATTTCTGACTGGGTGGAAGGCAAGAGAGCCCCGAAGCTT
ACATTTTATGTTAGCTGGTGGACTGACGCCAGAAAATGTTGGTGATGCGCTTAGATTAAATGGCGTTATTGGTGTGATG
TAAGCGGAGGTGTGGAGACAAATGGTGTAAAAGACTCTAACAAAATAGCAAATTTTCGTCAAAAATGCTAAGAAATAGGTT
ATTACTGAGTAGTATTTATTTAAGTATTGTTTGTGCACTTGCTGCAGGCC TTTTGAAAAGCAAGCATAAAAGATCTAAA
CATAAAAATCTGTAATAAACAAGATGTAAAGATAATGCTAAATCATTGGCTTTTGGATTGATTGTACAGGAAAATATAC
ATCGCAGGGGTTGACTTTTACCATTTACCCGATGGAATCAAACCTGTTtGAAGAG AATG
```

CGGGCCGC – NotI

GGCGGCC – Ascl

GGATCC – BamHI

GGTACC – KpnI

GAATTC – EcoRI

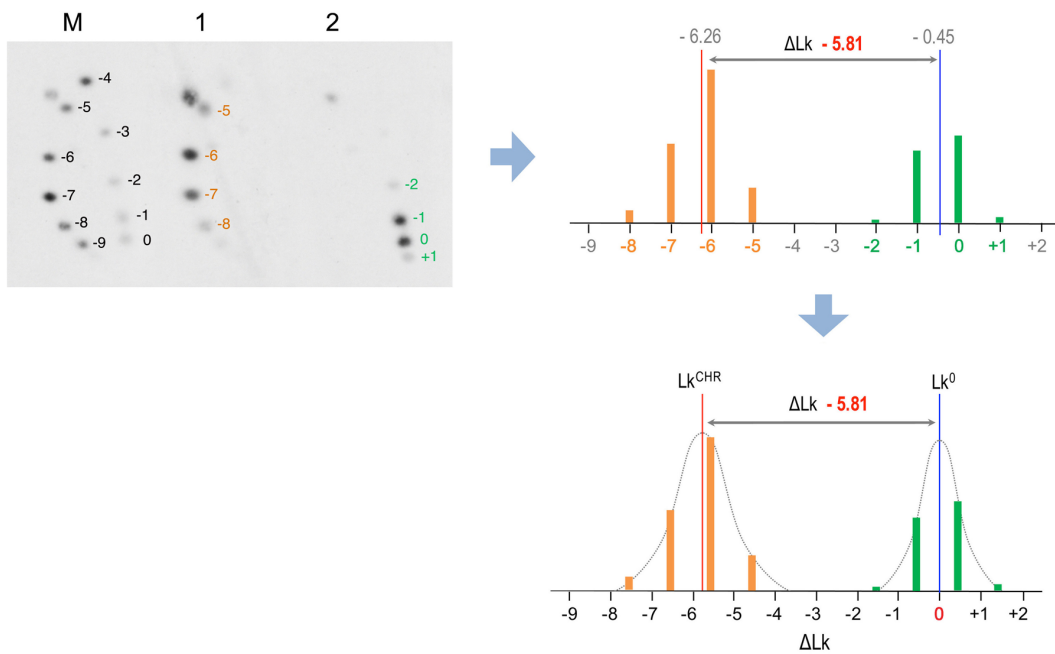
AAGCTT – HindIII

CEN2 elements: CDEIII CDEII CDEI

TRP1-ARS1 genomic fragment

TRP1 coding sequence - ARS1 sequence

Supplementary Fig 1. Construction of YCp1.3. A. PCR oligonucleotides used to amplify the TRP-ARS1 segment and the CEN2 of *S.cerevisiae*. B. Complete base pair sequence (5' to 3') of the YCp1.3 circle. Relevant restriction sites and the functional elements of YCp1.3 are color coded as indicated.



Supplementary Fig 2. Analysis of Lk distributions and calculation of ΔLk . Top-left: Lk markers (lane M) helped to count unambiguously the number of topoisomers between Lk distributions (lanes 1 and 2) resolved in the same 2D gel. An integer ΔLk value (colored) was initially assigned to each Lk topoisomer of the minichromosome (lane 1) and the relaxed DNA (lane 2). To this end, the most intense topoisomer of relaxed DNA was given the value $\Delta Lk=0$. Top-right: From the individual topoisomer intensities and assigned ΔLk values, the mean of each Lk distribution was calculated (grey). The ΔLk between the Lk distributions was then obtained by subtracting their means. Bottom: These results were illustrated in the figures by plotting the intensity of individual Lk topoisomers of minichromosomes and relaxed DNA along a scale of ΔLk units (x-axis), in which the value $\Delta Lk=0$ was corrected to the mean of the Lk distribution of the relaxed DNA (Lk^0). The mean of the Lk distribution of the minichromosome (Lk^{CHR}) was thereby ΔLk . This presentation of the data facilitated the comparison of distinct minichromosome clones, since all were aligned by Lk^0 .

A

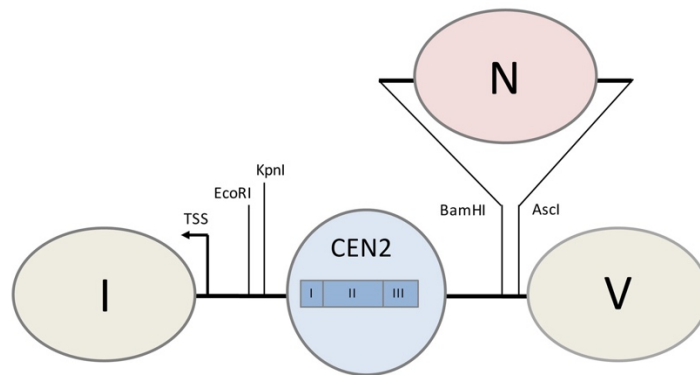
Adaptor with BamHI cohesive-end ligatable but not recutable

```
5'      *GGACGATTACAGCTACGTG      3'
3'      TCCTGCTAATGTCGATGCACCTAG  5'
```

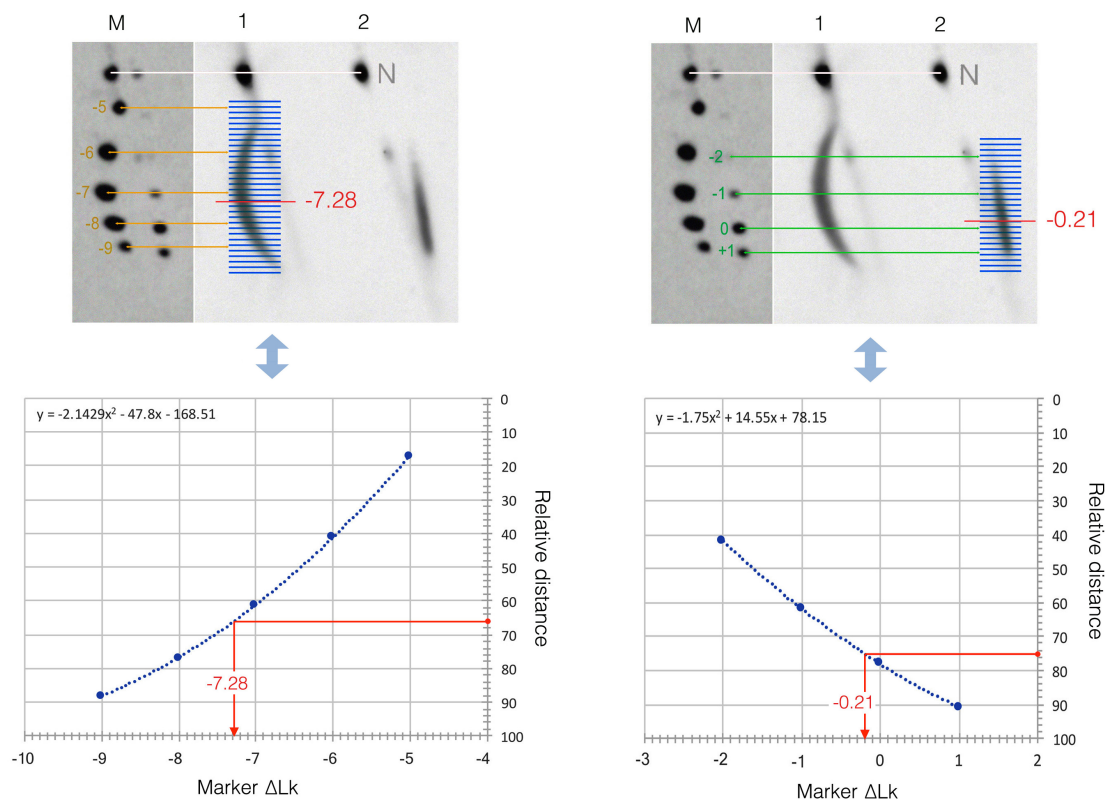
Adaptor with AscI cohesive-end ligatable but not recutable

```
5'      *CGTTCGGATCCGTTTAAACGTGAAGAGGTA ACTCAT      3'
3'      TGCAAGCCTAGGCAAATTTGCACTTCTCCATTGAGTAGCGC  5'
```

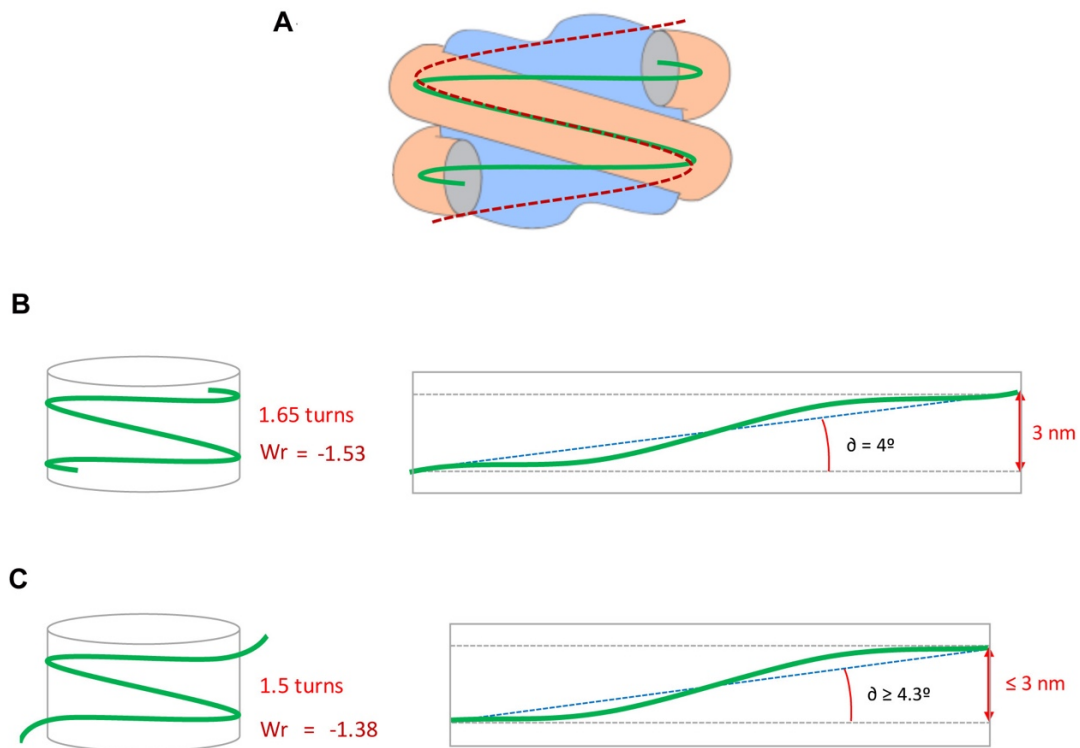
B



Supplementary Fig 3. Inclusion of a mononucleosomal library into YCp1.3. **A.** Adaptors ligated to A-tailed mononucleosomal DNA fragments. Note that the BamHI and AscI cohesive ends could be ligated but not recutted. This strategy facilitated their insertion in between the BamHI and AscI sites of YCp1.3. **B.** Position of the nucleosome library with respect to nucleosome V and the CEN2 elements of YCp1.3.



Supplementary Fig 4. Analysis of the pooled Lk distributions and calculation of ΔLk . An integer ΔLk value was assigned to individual Lk topoisomers of the marker (lane M, overexposed) as explained in Supplementary Figure 2. The signals of pooled Lk distributions of the minichromosomes (top left, lane 1) and of their relaxed DNAs (top right, lane 2) were quantified by bins (blue sections) and their mean was determined (red marks). In the bottom graphs, the gel position of the means (red dots) was interpolated with that of the Lk topoisomers of the marker (blue dots) to obtain the ΔLk value of the means in the marker frame. Y-axes are relative migration distances to the nicked DNA circles (N) in the first gel dimension. X-axes are the integer ΔLk values assigned to the marker. The ΔLk between the pool of minichromosomes and their relaxed DNAs was determined by subtracting the ΔLk values so obtained in the marker frame ($\Delta Lk = -7.28 - (-0.21) = -7.07$).



Supplementary Fig 5. Average pitch angle and effective writhe of nucleosomal DNA.

A. The turning of DNA around the histone core (green) does not follow a regular helical path (dashed red). The average pitch angle (δ) is then determined from the helical height achieved by the entire contour of wrapped DNA. The writhe is then given by $Wr = n(1 - \sin \delta)$, where n is the number of helical turns. **B.** If $n \approx 1.65$ (as in crystallized nucleosomes and chromatosomes), $Wr = -1.53$. **C.** If $n = 1.5$ (due to partial unwrapping or breathing motions of the core DNA), $Wr = -1.38$.

discusión general

DISCUSIÓN GENERAL

Los resultados de esta tesis demuestran que nuestra estrategia experimental para analizar el valor de ΔLk de minicromosomas circulares de pequeño tamaño permite determinar con precisión cómo los distintos elementos de la cromatina deforman el ADN. A diferencia de los estudios anteriores donde comparaban directamente distribuciones de Lk de distintos minicromosomas, nuestra aproximación determina el valor de ΔLk de cada minicromosoma usando el valor Lk_0 del ADN desnudo como referencia. Con este fin, obtenemos el valor de Lk_0 relajando los anillos de ADN a la misma temperatura que los minicromosomas *in vivo*. Posteriormente, comparamos el ΔLk de minicromosomas que contienen distintos insertos. Las diferencias de ΔLks ($\Delta\Delta Lk$) nos permiten calcular entonces con precisión la contribución individual de cada uno de los elementos insertados. Esta información estructural, capturada *in vivo*, difícilmente puede obtenerse mediante otras metodologías y es esencial para comprender la arquitectura de los mecanismos que gobiernan la biología del ADN.

Como prueba de la utilidad de esta estrategia, demostramos que el centrómero puntual de *S. cerevisiae* estabiliza un ΔLk de +0.6 unidades y que esta topología está determinada por los complejos proteicos unidos a CDEII y/o CDEIII, pero no a CDEI (Diaz-Ingelmo *et al.*, 2015). Para llegar a esta conclusión era necesario excluir otros mecanismos que podrían causar las diferencias de Lk entre minicromosomas con y sin centrómero. En primer lugar, descartamos alteraciones de los valores de Lk durante la disrupción celular y la extracción de ADN. Para ello, fijamos los valores de ΔLk *in vivo* mediante una disolución fría (-20°C) compuesta por etanol y tolueno que inactiva de manera irreversible las topoisomerasas celulares. En segundo lugar, se descarta que los valores de ΔLk pudieran deberse a diferencias de tensión helicoidal (relajable). Con esa finalidad, los minicromosomas se solubilizan e incuban con topoisomerasas para corroborar que los valores de ΔLk fijados *in vivo* no cambiaban al relajar los minicromosomas *in vitro*. En tercer lugar, descartamos que los cambios en el ΔLk producidos por el centrómero puntual se deban a alteraciones de la topología del ADN inducidas en las regiones vecinas. Para ello,

demostramos que, al reducir el tamaño de los minicromosomas y al insertar nuevas secuencias, no hay pérdida de nucleosomas vecinos y se mantienen las funciones del gen *TRP1* y de la región *ARS1*.

Nuestra estrategia nos ha permitido comparar los centrómeros puntuales de distintos cromosomas de *S. cerevisiae* y la contribución de cada CDE en la topología del complejo. Por ejemplo, en todos los centrómeros analizados, se encontró la estabilización de +0.6 unidades de *Lk*. Al ser secuencias de diferente longitud, concluimos que este valor es independiente de la longitud variable de CDEII y, por lo tanto, de la fase de rotación entre CDEI y CDEIII. Al observar que las mutaciones que evitan la unión de Cbf1 a CDEI no alteran la topología del ADN, dedujimos la ausencia de un puente proteico entre CDEI y CDEIII. Por el contrario, demostramos que la estabilización de +0.6 unidades de *Lk* depende de la integridad de CDEII y CDEIII. Para ello, los complejos de proteínas que interactúan con CDEII y CDEIII deben deformar el ADN en forma de ΔTw y/o ΔWr , tal que la suma de ΔTw y ΔWr es +0.6. Este escenario sugiere que un trayecto levógiro del ADN ($\Delta Wr < 0$) es muy improbable. Si fuera así, CDEII y CDEIII tendrían que estabilizar un $\Delta Tw > +0.6$, lo que implicaría un inverosímil valor h de ≈ 9 pb en lugar de 10.5 pb. El mecanismo más simple para estabilizar +0.6 unidades de *Lk* es un giro dextrógiro del ADN sobre el complejo de histonas ($\Delta Wr > 0$). Este giro dextrógiro podría estar inducido en CDEII mediante el complejo de histonas con la variante CENP-A. Sin embargo, no hay evidencia de complejos de histonas inherentemente dextrógiros, aunque es conocido que algunas partículas sub-octaméricas (tetrasomas o hemisomas) son más flexibles que el octámero para cambiar su quiralidad en función del estado de superenrollamiento del ADN (Hamiche *et al.*, 1996; Vlijm *et al.*, 2015). Bajo estas premisas, concluimos que el giro dextrógiro del ADN debería estar inducido por el complejo CBF3.

El ensamblaje del complejo CBF3 a la región centromérica es un proceso altamente regulado. CBF3 reconoce la secuencia CDEIII a través de Cep3 (Espelin *et al.*, 1997; Pietrasanta *et al.*, 1999; Russell *et al.*, 1999; Purvis & Singleton, 2008). En contraste, Ndc10 no reconoce una secuencia de ADN específica, sino que interactúa con CDEIII y con

regiones ricas en AT, como las presentes en CDEII (Espelin *et al.*, 1997; Espelin *et al.*, 2003). Posteriores estudios estructurales revelaron que el dímero Ndc10 se une a dos segmentos separados de ADN, cada uno de aproximadamente 10 pb (Cho & Harrison, 2012). Por lo tanto, mientras una subunidad de Ndc10 se une a CDEIII junto a Cep3, la otra subunidad puede interactuar con un segmento de ADN más alejado, generando de este modo un bucle o *loop*. Curiosamente, esta capacidad del complejo CBF3 para formar un bucle de ADN ya se había podido observar mediante microscopía de fuerza atómica (AFM) anteriormente (Pietrasanta *et al.*, 1999) (Figura D1a). Las imágenes revelaron que CBF3 acortaba la longitud del ADN unas 70 pb y generaba un ángulo de 55° entre los segmentos de entrada y salida del ADN. Con estos datos y conociendo la geometría del dímero de Ndc10 unido a dos segmentos de ADN (Cho & Harrison, 2012), calculamos la trayectoria del supuesto bucle de ADN y el valor teórico de ΔWr . El resultado fue un valor de $\Delta Wr = +0.63$ unidades, el cual prácticamente coincidía con el valor experimental de $\Delta Lk +0.6$ obtenido en nuestros análisis de minicromosomas *in vivo*. A partir de estas premisas, propusimos un modelo del centrómero puntual donde el dímero Ndc10 genera un bucle dextrógiro de ADN al anclar un segmento de entrada en CDEIII (al lado de Cep3) y un segmento de salida en CDEII (Figura D1b). En este bucle, de entre 70 y 80 pb, es donde, posteriormente, se acomoda el complejo de histonas CENP-A, reclutado a través de Scm3 (una chaperona CENP-A reconocida por Ndc10) (Cho & Harrison, 2011; Wisniewski *et al.*, 2014). Nuestro modelo explica por qué la interacción de las histonas con CDEII puede ocurrir en registros de pares de bases variables (dos orientaciones y dos marcos helicoidales posibles) (Henikoff *et al.*, 2014). Dado que CDEI se encuentra fuera del bucle de ADN, nuestro modelo también explica por qué las interacciones de Cbf1 con CBF3 no son relevantes para determinar la topología del ADN en el centrómero puntual. Finalmente, cuatro años después de la publicación de nuestro trabajo (Diaz-Ingelmo *et al.*, 2015), una primera estructura del centrómero puntual de *S. cerevisiae* se obtuvo mediante crio-microscopía (Cryo-EM) (Zhang *et al.*, 2019) (Figura D1c). Esta estructura confirma nuestro modelo construido mediante análisis topológicos del ADN.

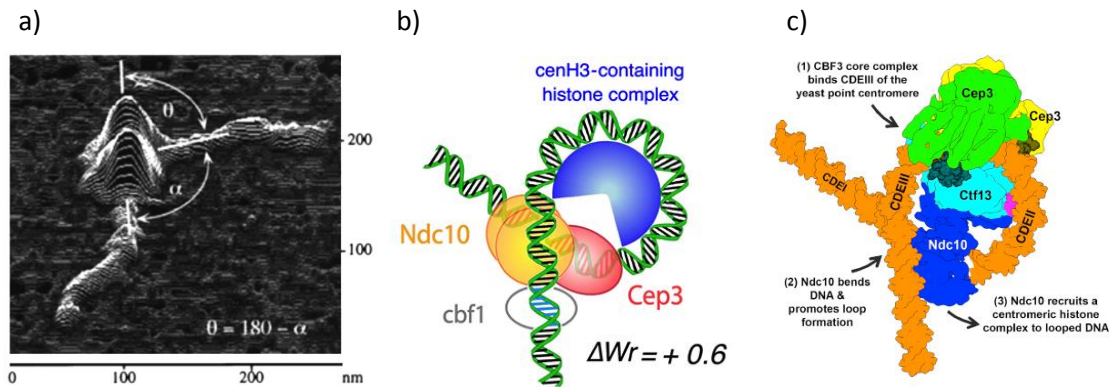


Figura D1. Modelos estructurales del centrómero de *S. cerevisiae*. (a) Imagen de microscopía de fuerza atómica del complejo CBF3 unido al ADN (Pietrasanta *et al.*, 1999). (b) Modelo presentado a partir de nuestro análisis topológico del centrómero puntual de *S. cerevisiae* (Díaz-Ingelmo *et al.*, 2015). (c) Estructura observada mediante Cryo-EM (Zhang *et al.*, 2019).

El análisis topológico de minicromosomas circulares de *S. cerevisiae* con y sin centrómero también nos reveló que el Lk estabilizado por los nucleosomas adyacentes era más negativo que el valor de $\Delta Lk -1$ por nucleosoma asumido. Esta observación fue la que nos motivó a llevar a cabo una revisión de la paradoja del número de enlace, lo que constituyó el objetivo principal de la presente tesis. Siguiendo la misma estrategia utilizada con el centrómero puntual, construimos e insertamos una librería de nucleosomas en los pequeños minicromosomas y analizamos su topología. Los resultados revelaron que la mayoría de los nucleosomas nativos estabilizan un ΔLk cercano a -1.26 (Segura *et al.*, 2018).

La discrepancia entre nuestro valor (-1.26) y el valor previamente establecido (-1.0), puede deberse a los distintos modelos de cromatina y la imprecisión de los métodos que han sido utilizados para el cálculo del ΔLk estabilizado por los nucleosomas. En primer lugar, una fuente de inexactitud radica en determinar el número de nucleosomas ensamblados en las moléculas de ADN circulares. Los estudios previos, con el minicromosoma SV40 como modelo de cromatina, utilizaban la microscopía electrónica para contar el número de nucleosomas o partículas similares a nucleosomas ensamblados.

Los números obtenidos variaban entre 20 – 27 nucleosomas según experimento y laboratorio (Germond *et al.*, 1975; Germond, 1978; Saragosti *et al.*, 1980; Sogo *et al.*, 1986). Sin embargo, el mapeo más reciente de la cromatina de SV40 ha revelado que esta variabilidad no sólo es experimental. Tanto los minicromosomas intracelulares como los viriones de SV40 pueden presentar un número de variable de nucleosomas junto con modificaciones epigenéticas que alterarían la organización del nucleosoma en función de la etapa de infección en la que se encuentra (Kumar *et al.*, 2017). Por lo tanto, esta variabilidad en el número de nucleosomas genera valores de ΔLk imprecisos, especialmente cuando el conteo de nucleosomas y los análisis de topología de ADN se realizaron con muestras independientes o en distintos laboratorios (Keller, 1975; Shure & Vinograd, 1976). Como resultado, el modelo SV40 generaba una amplia gama de valores de ΔLk , incluido el valor -1.25 por nucleosoma (Shure & Vinograd, 1976). De la misma manera, a partir de los estudios de reconstrucción de la cromatina *in vitro* donde utilizaban la microscopía electrónica para el conteo de nucleosomas también obtenían valores imprecisos (Simpson *et al.*, 1985; Norton *et al.*, 1989). Además, el uso de secuencias repetitivas de ADN en tándem para la reconstrucción de la cromatina genera una regularidad que puede desviar notablemente los valores de ΔLk por nucleosoma. Estos valores varían entre -0.8 y -1.4 unidades dependiendo de la longitud periódica del ADN *linker* entre secuencias de posicionamiento (Nikitina *et al.*, 2017).

Por otro lado, la segunda gran fuente de imprecisión en estudios anteriores fue el cálculo de ΔLk a partir de las bandas de ADN observadas obtenidas a partir de electroforesis en geles de agarosa. En la mayoría de los estudios sobre SV40 y cromatina reconstituida, se asignaba erróneamente la posición del centro de las distribuciones de Lk (Lk_0) al topoisómero Lk más lento o al más abundante (Keller *et al.*, 1975; Germond *et al.*, 1975; Germond *et al.*, 1978; Simpson *et al.*, 1985; Norton *et al.*, 1989). De la misma manera, en las mediciones de ΔLk utilizando minicromosomas circulares de levadura como modelo, se confundía la posición del ADN relajado con la del ADN covalentemente abierto (con *nicks*). Como, por ejemplo, en el caso del minicromosomas *TRP1ARS1*, el cual contiene 7

nucleosomas, y al que se asignó un valor de ΔLk de -7 (Pederson *et al.*, 1986). Un valor que contrasta notoriamente con el valor obtenido en nuestro estudio, $\Delta Lk = -9.6$ unidades.

Finalmente, en la mayoría de los estudios anteriores no quedaba claro si el procesamiento de las muestras de cromatina y/o la relajación del ADN se realizaban bajo la misma temperatura. Dado que el valor h (*helical repeat*) del ADN disminuye a medida que la temperatura disminuye (Depew & Wang, 1975), un ADN relajado a 4°C produce valores de Lk hasta 1 unidad por kb más altos que cuando se relaja a 37°C, tal como se demuestra en esta memoria (Diaz-Ingelmo *et al.*, 2015, Figura S2).

Mediante la estrategia experimental presentada en esta tesis eliminamos todas las anteriores incertidumbres en el número de nucleosomas y en el cálculo de ΔLk . Los minicromosomas utilizados como modelo para calcular el ΔLk nucleosomal presentan 5 nucleosomas con posiciones muy bien definidas, y 3 elementos de cromatina específicos (promotor TRP1, ARS1 y CEN2). Su pequeño tamaño también permite descartar posibles efectos debidos a un plegamiento de orden superior de la cromatina. Nuestros resultados experimentales corroboran el valor de ΔLk estabilizado por nucleosomas individuales de dos maneras. Primero, al promediar el ΔLk de los minicromosomas por su número de nucleosomas, obtenemos un ΔLk de -1.28 por nucleosoma. Sin embargo, esta medición no permite descartar una posible variabilidad debida a otros elementos estructurales distintos de los nucleosomas. Este problema se elimina tras la inserción de la librería de nucleosomas y determinando la ganancia de ΔLk ($\Delta\Delta Lk$) producida por el nuevo elemento.

Aparte de obtener el valor ΔLk de -1.26 por nucleosoma, también observamos muy poca dispersión en el ΔLk restringido por la librería de nucleosomas. Esta observación nos indica que la mayoría de los nucleosomas estabilizan una topología de ADN similar. Si fuera cierto que los nucleosomas nativos estabilizan cada uno un $\Delta Lk = -1.0$, nuestros resultado de -1.26 difícilmente podrían explicarse. Sólo podría tener lugar si los minicromosomas hubiesen contenido más nucleosomas de los que físicamente caben en la longitud del ADN, lo que queda descartado mediante los resultados de digestión con MNasa. Por otro

lado, un ΔLk de -1.0 por nucleosoma podría ser cierto si los minicromosomas se hubieran compactado espacialmente mediante la adopción de un fuerte quiralidad negativa, es decir, un $\Delta Wr \approx -1.0$. Tal plegamiento implicaría que las longitudes del ADN *linker* y las consiguientes orientaciones rotacionales entre nucleosomas adyacentes tendrían que ser muy parecidas en todos los minicromosomas. Sin embargo, la librería de nucleosomas insertada comprende segmentos de varias longitudes (156 ± 8 pb) y, aun así, los minicromosomas resultantes estabilizaban valores de ΔLk muy similares.

Tras nuestros resultados, el valor obtenido de $\Delta Lk = -1.26$ nos lleva a revisar la paradoja del Lk del ADN nucleosomal en relación a la ecuación general, $\Delta Lk = \Delta Tw + \Delta Wr$. Teniendo en cuenta que las 147 pb de ADN del *core* nucleosomal presentan un valor ΔTw de +0.20 unidades (White & Bauer, 1989; Bates & Maxwell, 2005), la estabilización de -1.26 unidades de ΔLk implica que la deformación en forma de ΔWr debería ser de aproximadamente -1.46. Nosotros, según las dimensiones cristalográficas del nucleosoma, hemos calculado el valor de ΔWr para diferentes grados de enrollamiento del ADN nucleosomal. Para ello utilizamos la ecuación del Wr de una curva helicoidal, $Wr = n(1 - \sin \delta)$, donde n es el número de vueltas helicoidales y δ es el ángulo de cabeceo (*helical pitch*) de dichas vueltas (Vologodskii & Cozzarelli, 1994). Así obtenemos un valor de $Wr = -1.53$ cuando el ADN del *core* completa 1.65 vueltas levóginas alrededor del octámero de histonas. Esta conformación correspondería a la estructura típica del nucleosoma canónico que ha sido cristalizada (Luger *et al.*, 1997). La conformación de 1.65 vueltas levóginas equivaldría también al de los cromatosomas, donde la histona H1 interacciona y fija los segmentos de entrada y salida del ADN cuando se cruzan en un ángulo de aproximadamente 60 grados (Bednar *et al.*, 2017). Por lo tanto, es probable que el valor de $Wr = -1.53$ refleje el límite superior del Wr del ADN del *core* nucleosomal en solución. Sin embargo, *S. cerevisiae* carece de histona H1 y presenta pocas proteínas asociadas al ADN *linker* (Panday & Grove, 2017). En este caso, estudios *in vitro* demuestran que la dinámica conformacional de los nucleosomas en disolución implica desenrollamientos parciales y transitorios del ADN (*breathing*) (Tomschik *et al.*, 2005; Li *et al.*, 2005; Bucceri

et al., 2006). Estas fluctuaciones generan una reducción del valor de Wr del ADN nucleosomal y, por lo tanto, una reducción de su valor promedio. Nosotros calculamos que al reducir el giro del ADN de 1.65 a 1.5 vueltas levóginas, el valor Wr se reduce a -1.38. Por lo tanto, un valor promedio del ΔWr cercano a -1.46 por nucleosoma es realista. Este valor, junto con $\Delta Tw = +0.2$ y $\Delta Lk = -1.26$, resuelve la paradoja del número de enlace del ADN nucleosomal.

La estrategia experimental y los resultados obtenidos en esta tesis contribuyen a una mejor comprensión de cómo distintos complejos proteicos de la cromatina deforman la doble hélice del ADN, condición fundamental de la que deriva la regulación genómica y la arquitectura cromosómica. En el caso de los nucleosomas, nuestros resultados permiten entender mejor cómo los complejos ADN-histona confinan y/o amortiguan el superenrollamiento del ADN durante los procesos de actividad génica. En el caso de complejos singulares, como sería el centrómero puntual, conocer la topología del ADN permite entender mejor los mecanismos moleculares y fuerzas ejercidas sobre el ADN durante funciones especializadas de la cromatina.

Por último, nuestra estrategia para examinar la topología de librerías de elementos de la cromatina, en lugar de examinarlos de forma individual, abre la posibilidad de construir el primer catálogo de elementos de la cromatina en función de cómo deforman el ADN. Esta idea, cuyo desarrollo es incipiente y no se ha incluido en esta tesis, se basa en la posibilidad de realizar secuenciación masiva del ADN extraído de distintas secciones de un gel de agarosa, en el que se han resuelto las distribuciones de Lk de minicromosomas que contienen una librería de insertos. Como muestra la Figura D2, cada elemento de la librería producirá una distribución de Lk específica. Pero al estar miles de distribuciones Lk superpuestas en una única calle del gel, no es posible diferenciarlas. Sin embargo, secuenciando el ADN presente en distintas secciones del gel y anotando la abundancia relativa de cada secuencia, es posible reconstruir la distribución de Lk asociada a cada

secuencia y, por tanto, calcular su ΔLk . Esta aproximación permitiría calcular el valor de ΔLk estabilizado por miles de elementos de la cromatina simultáneamente, permitiendo descubrir complejos con topologías singulares y proporcionando una inédita perspectiva de la arquitectura del genoma.

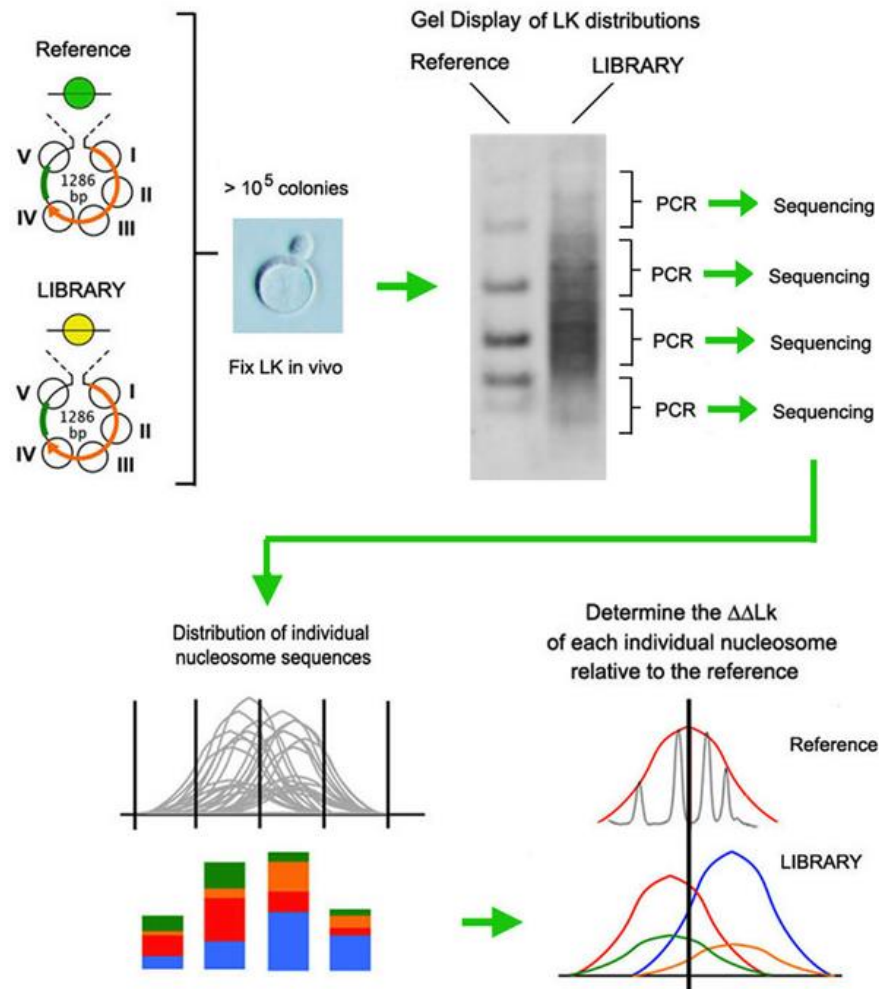


Figura D2. **Diseño experimental para la construcción de un catálogo genómico del ΔLk estabilizado por elementos de la cromatina.** La amplificación mediante PCR y posterior secuenciación del ADN extraído de distintas posiciones del gel mediante NGS (*Next Generation Sequencing*), permite reconstruir la distribución Lk asociada a cada una de las secuencias que constituyen la librería de nucleosomas.

conclusiones

1. El uso de minicromosomas circulares permite analizar *in vivo* el valor de ΔLk estabilizado por los elementos estructurales de la cromatina de *Saccharomyces cerevisiae* de manera precisa.
2. Los nucleosomas nativos de *S. cerevisiae* estabilizan un valor medio de $\Delta Lk = -1.26$.
3. Teniendo en cuenta los valores promedios y revisados de $\Delta Wr = -1.46$ y $\Delta Tw = +0.2$ en el ADN nucleosomal, el valor de $\Delta Lk = -1.26$ resuelve la llamada paradoja del número de enlace.

bibliografía

- Adrian, M., ten Heggeler-Bordier, B., Wahli, W., Stasiak, A., Stasiak, A., & Dubochet, J. (1990). Direct visualization of supercoiled DNA molecules in solution. *The EMBO Journal*, *9*(13), 4551-4554.
- Armeev, G., Gribkova, A., Pospelova, I., Komarova, G., & Shaytan, A. (2019). Linking chromatin composition and structural dynamics at the nucleosome level. *Current Opinion in Structural Biology*, *56*, 46-55.
- Arsuaga, J., Vázquez, M., Trigueros, S., Sumners, D., & Roca, J. (2002). Knotting probability of DNA molecules confined in restricted volumes: DNA knotting in phage capsids. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(8), 5337-5377.
- Bancaud, A., Conde e Silva, N., Barbi, M., Wagner, G., Allemand, J.F., Mozziconacci, J., Lavelle, C., Croquette, V., Victor, J., Prunell, A., & Viovy, J. L. (2006). Structural plasticity of single chromatin fibers revealed by torsional manipulation. *Nature Structural & Molecular Biology*, *13*(5), 444-450.
- Barski, A., Cuddapah, S., Cui, K., Roh, T., Schones, D., Wang, Z., Wei, G., Chepelev, I., & Zhao, K. (2007). High-Resolution Profiling of Histone Methylations in the Human Genome. *Cell*, *129*(4), 823-837.
- Bates, A., & Maxwell, A. (2005). *DNA Topology* (2nd edition ed.). Oxford University Press.
- Bauer, W. (1978). Structure and Reactions of Closed Duplex DNA. *Annual Review of Biophysics and Bioengineering*, *7*(1), 287-313.
- Baxevanis, A., Godfrey, J., & Moudrianakis, E. (1991). Associative Behavior of the Histone (H3-H4)₂ Tetramer: Dependence on Ionic Environment. *Biochemistry*, *30*(36), 8817-8823.
- Becker, P., & Workman, J. (2013). Nucleosome remodeling and epigenetics. *Cold Spring Harbor Perspectives in Biology*, *5*(9).
- Bednar, J., Garcia-Saez, I., Boopathi, R., Cutter, A., Papai, G., Reymer, A., Syed, S., Lone, I., Tonchev, O., Crucifix, C., Menoni, H., Papin, C., Skoufias, D., Kurumizaka, H., Lavery, R., Hamiche, A., Hayes, J., Schultz, P., Angelov, D., Petosa, C., & Dimitrov, S. (2017). Structure and Dynamics of a 197 bp Nucleosome in Complex with Linker Histone H1. *Molecular Cell*, *66*(3), 384-397.e8.
- Bloom, K., & Carbon, J. (1982). Yeast centromere DNA is in a unique and highly ordered structure in chromosomes and small circular minichromosomes. *Cell*, *29*(2), 305-317.
- Bloom, K., Amaya, E., Carbon, J., Clarke, L., Hill, A., & Yeh, E. (1984). Chromatin conformation of yeast centromeres. *Journal of Cell Biology*, *99*(5), 1559-1568.
- Bucceri, A., Kapitzka, K., & Thoma, F. (2006). Rapid accessibility of nucleosomal DNA in yeast on a second time scale. *EMBO Journal*, *25*(13), 3123-3132.

- Champoux, J. (1990). 6 Mechanistic Aspects of Type-I Topoisomerases. *Cold Spring Harbor Monograph Archive*, 20(0), 217-242.
- Cho, U., & Harrison, S. (2011). Recognition of the centromere-specific histone Cse4 by the chaperone Scm3. *Proceedings of the National Academy of Sciences of the United States of America*, 108(23), 9367-9371.
- Cho, U., & Harrison, S. (2012). Ndc10 is a platform for inner kinetochore assembly in budding yeast. *Nature Structural and Molecular Biology*, 19(1), 48-56.
- Clarke, L. (1998). Centromeres: Proteins, protein complexes, and repeated domains at centromeres of simple eukaryotes. *Current Opinion in Genetics and Development*, 8(2), 212-218.
- Corbett, K., & Berger, J. (2001). Structure, molecular mechanisms and evolutionary relationships in DNA topoisomerases. *Annual Review of Biophysics and Biomolecular Structure*, 33, 95-118.
- Corbett, K., Shultzaberger, R., & Berger, J. (2004). The C-terminal domain of DNA gyrase A adopts a DNA-bending β -pinwheel fold. *Proceedings of the National Academy of Sciences of the United States of America*, 101(19), 7293-7298.
- Davey, C., Sargent, D., Luger, K., Maeder, A., & Richmond, T. (2002). Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *Journal of Molecular Biology*, 319(5), 1097-1113.
- De Lucia, F., Alilat, M., Sivolob, A., & Prunell, A. (1999). Nucleosome Dynamics III. Histone tail-dependent fluctuation of nucleosomes between open and closed DNA conformations. Implications for chromatin dynamics and the linking number paradox. A relaxation study of mononucleosomes on DNA minicircles. *Journal of Molecular Biology*, 285(3), 1101-1119.
- Depew, R., & Wang, J. (1975). Conformational fluctuations of DNA helix. *Proceedings of the National Academy of Sciences of the United States of America*, 72(11), 4275-4279.
- Díaz-Ingelmo, O., Martínez-García, B., Segura, J., Valdés, A., & Roca, J. (2015). DNA Topology and Global Architecture of Point Centromeres. *Cell Reports*, 13(4), 667-677.
- Dickerson, R., Drew, H., Conner, B., Wing, R., Fratini, A., & Kopka, M. (1982). The anatomy of A-, B-, and Z-DNA. *Science*, 216(4545), 475.
- Dorigo, B., Schalch, T., Kulangara, A., Duda, S., Schroeder, R., & Richmond, T. (2004). Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science (New York, N.Y.)*, 306(5701), 1571-3.
- Draizen, E.J., Shaytan, A.K., Marino-Ramirez, L., Talbert, P.B., Landsman, D., & Panchenko, A.R. (2016). Histone DB 2.0: a histone database with variants- an integrated resource to explore histones and their variants. *Database*. PMID: 26989147.

- Eng, W., Pandit, S., & Sternglanz, R. (1989). Mapping of the active site tyrosine of eukaryotic DNA topoisomerase I. *The Journal of biological chemistry*, 264(23), 13373-6.
- Espelin, C., Kaplan, K., & Sorger, P. (1997). Probing the architecture of a simple kinetochore using DNA-protein crosslinking. *Journal of Cell Biology*, 139(6), 1383-1396.
- Espelin, C., Simons, K., Harrison, S., & Sorger, P. (2003). Binding of the Essential *Saccharomyces cerevisiae* Kinetochore Protein Ndc10p to CDEII. *Molecular Biology of the Cell*, 14(11), 4557-4568.
- Finch, J., Lutter, L., Rhodes, D., Brown, R., Rushton, B., Levitt, M., & Klug, A. (1977). Structure of nucleosome core particles of chromatin. *Nature*, 269(5623), 29-36.
- Fragoso, G., John, S., Roberts, M., & Hager, G. (1995). Nucleosome positioning on the MMTV LTR results from the frequency-biased occupancy of multiple frames. *Genes and Development*, 9(15), 1933-1947.
- Franklin, R., & Gosling, R. (1953). Molecular Configuration in Sodium Thymonucleate. *Nature*, 171(4356), 740-741.
- Fuller, F. (1978). Decomposition of the linking number of a closed ribbon: A problem from molecular biology. *Proceedings of the National Academy of Sciences of the United States of America*, 75(8), 3557-3561.
- Furuyama, T., & Henikoff, S. (2009). Centromeric nucleosomes induce positive DNA supercoils. *Cell*, 138(1), 104-113.
- García, A., González, S., & Antequera, F. (2017). Nucleosomal organization and DNA base composition patterns. *Nucleus*, 8(5), 469-474.
- Gellert, M., Mizuuchi, K., O'Dea, M., & Nash, H. (1976). DNA gyrase: an enzyme that introduces superhelical turns into DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 73(11), 3872-3876.
- Germond, J., Hirt, B., Oudet, P., Gross-Bellard, M., & Chambon, P. (1975). Folding of the DNA double helix in chromatin-like structures from simian virus 40. *Proceedings of the National Academy of Sciences*, 72(5), 1843-1847.
- Germond, J. (1978). Simian virus 40 DNA extracted from infected cells with sodium deoxycholate no longer reflects its in vivo superhelix density. *Journal of virology*, 26(3), 681-5.
- Hamiche, A., Carot, V., Alilat, M., De Lucia, F., O'Donohue, M., Révet, B., & Prunell, A. (1996). Interaction of the histone (H3-H4)₂ tetramer of the nucleosome with positively supercoiled DNA minicircles: Potential flipping of the protein from a left- to a right-handed superhelical form. *Proceedings of the National Academy of Sciences of the United States of America*, 93(15), 7588-7593.

- Hammond, C., Strømme, C., Huang, H., Patel, D., & Groth, A. (2017). Histone chaperone networks shaping chromatin function. *Nature Reviews Molecular Cell Biology*, *18*(3), 141-158. Nature Publishing Group.
- Hanai, R., & Roca, J. (1999). Two-dimensional agarose-gel electrophoresis of DNA topoisomers. *Methods in molecular biology (Clifton, N.J.)*, *94*, 19-27.
- Hayes, J., Clark, D., & Wolffe, A. (1991). Histone contributions to the structure of DNA in the nucleosome. *Proceedings of the National Academy of Sciences of the United States of America*, *88*. 6829-6833.
- Henikoff, S., Ramachandran, S., Krassovsky, K., Bryson, T., Codomo, C., Brogaard, K., Widom, J., Wang, J., & Henikoff, J. (2014). The budding yeast centromere DNA element II wraps a stable Cse4 hemisome in either orientation in vivo. *eLife*, *2014*(3).
- Henikoff, S., & Smith, M. (2015). Histone variants and epigenetics. *Cold Spring Harbor Perspectives in Biology*, *7*(1), a019364.
- Hsieh, T.S. (1990). 7 Mechanistic Aspects of Type-II DNA Topoisomerases. *Cold Spring Harbor Monograph Archive*, *20*(0), 243-263.
- Huang, C., Chang, K., Cui, H., & Jayaram, M. (2011). Histone H3-variant Cse4-induced positive DNA supercoiling in the yeast plasmid has implications for a plasmid origin of a chromosome centromere. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(33), 13671-13676.
- Iyer, V., & Struhl, K. (1995). Poly(dA:dT), a ubiquitous promoter element that stimulates transcription via its intrinsic DNA structure. *The EMBO Journal*, *14*(11), 2570-2579.
- Jiang, C., & Pugh, B. (2009). Nucleosome positioning and gene regulation. Advances through genomics. *Nature Reviews Genetics*, *10*(3), 161-172.
- Kebede, A., Schneider, R., & Daujat, S. (2015). Novel types and sites of histone modifications emerge as players in the transcriptional regulation contest. *FEBS Journal*, *282*(9), 1658-1674.
- Keller, W. (1975). *Determination of the number of superhelical turns in simian virus 40 DNA by gel electrophoresis (DNA-relaxing enzyme/ethidium bromide/helix unwinding angle/chromatin structure)*.
- Kireeva, M., Walter, W., Tchernajenko, V., Bondarenko, V., Kashlev, M., & Studitsky, V. (2002). Nucleosome remodeling induced by RNA polymerase II: Loss of the H2A/H2B dimer during transcription. *Molecular Cell*, *9*(3), 541-552.

- Kumar, M., Christensen, K., Woods, B., Dettlaff, A., Perley, D., Scheidegger, A., Balakrishnan, L., & Milavetz, B. (2017). Nucleosome positioning in the regulatory region of SV40 chromatin correlates with the activation and repression of early and late transcription during infection. *Virology*, *503*, 62-69.
- Lai, W., & Pugh, B. (2017). Understanding nucleosome dynamics and their links to gene expression and DNA replication. *Nature reviews. Molecular cell biology*, *18*(9), 548-562.
- Li, G., & Widom, J. (2004). Nucleosomes facilitate their own invasion. *Nature Structural and Molecular Biology*, *11*(8), 763-769.
- Li, G., Levitus, M., Bustamante, C., & Widom, J. (2005). Rapid spontaneous accessibility of nucleosomal DNA. *Nature Structural and Molecular Biology*, *12*(1), 46-53.
- Liu, L., Perkocha, L., Calendar, R., & Wang, J. (1981). Knotted DNA from bacteriophage capsids. *Proceedings of the National Academy of Sciences of the United States of America*, *78*(9 II), 5498-5502.
- López-García, P., & Forterre, P. (1997). DNA topology in hyperthermophilic archaea: Reference states and their variation with growth phase, growth temperature, and temperature stresses. *Molecular Microbiology*, *23*, 1267-1279.
- Luger, K., Mäder, A., Richmond, R., Sargent, D., & Richmond, T. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*, *389*, 251-260.
- Luger, K., Dechassa, M. L., & Tremethick, D. J. (2012). New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? *Nature reviews. Molecular cell biology*, *13*(7), 436-447.
- Lynn, R. M., Bjornsti, M. A., Caron, P. R., & Wang, J. C. (1989). Peptide sequencing and site-directed mutagenesis identify tyrosine-727 as the active site tyrosine of *Saccharomyces cerevisiae* DNA topoisomerase I. *Proceedings of the National Academy of Sciences of the United States of America*, *86*(10), 3559-3563.
- Maeshima, K., Hihara, S., & Eltsov, M. (2010). Structure of the “30 nm” chromatin fibre: a key role for the linker histone. *Current Opinion in Cell Biology*, *22*(3), 291-297.
- Malik, H., & Henikoff, S. (2003). Phylogenomics of the nucleosome. *Nature structural biology*, *10*(11), 882-91.
- Malik, H., & Henikoff, S. (2009). Major evolutionary transitions in centromere complexity. *Cell*, *138*(6), 1067-1082.
- Mariño-Ramírez, L., King, I., & Landsman, D. (2006). Multiple independent evolutionary solutions to core histone gene regulation. *Genome Biology*, *7*(12): R122.

- Martínez-García, B., Fernández, X., Díaz-Ingelmo, O., Rodríguez-Campos, A., Manichanh, C., & Roca, J. (2014). Topoisomerase II minimizes DNA entanglements by proofreading DNA topology after DNA strand passage. *Nucleic Acids Research*, *42*(3), 1821-1830.
- Mavrich, T., Jiang, C., Ioshikhes, I., Li, X., Venters, B., Zanton, S., Tomsho, L., Qi, J., Glase, R., Schuster, S., Gilmour, D., Albert, I., & Pugh, B. (2008a). Nucleosome organization in the *Drosophila* genome. *Nature*, *453*(7193), 358-362.
- Mavrich, T., Ioshikhes, I., Venters, B., Jiang, C., Tomsho, L., Qi, J., Schuster, S., Albert, I., & Pugh, B. (2008b). A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome Research*, *18*(7), 1073-1083.
- McLeod, S., & Johnson, R. (2001). Control of transcription by nucleoid proteins. *Current Opinion in Microbiology*, *4*(2), 152-159.
- Nikitina, T., Norouzi, D., Grigoryev, S., & Zhurkin, V. (2017). DNA topology in chromatin is defined by nucleosome spacing. *Science Advances*, *3*(10), e1700957.
- Noll, M. (1974). Internal structure of the chromatin subunit. *Nucleic Acids Res.*, *1*, 1573-1578.
- Norton, V., Imai, B., Yau, P., & Bradbury, E. (1989). Histone acetylation reduces nucleosome core particle linking number change. *Cell*, *57*(3), 449-457.
- Olins, A., & Olins, D. (1974). Spheroid chromatin units (v bodies). *Science*, *183*(4122), 330-332.
- Panday, A., & Grove, A. (2017). Yeast HMO1: Linker Histone Reinvented. *Microbiology and Molecular Biology Reviews*, *81*(1).
- Papillon, J., Ménétret, J., Batisse, C., Hélye, R., Schultz, P., Potier, N., & Lamour, V. (2013). Structural insight into negative DNA supercoiling by DNA gyrase, a bacterial type 2A DNA topoisomerase. *Nucleic Acid Research*, *41*(16), 7815-7827.
- Partensky, P., & Narlikar, G. (2009). Chromatin Remodelers Act Globally, Sequence Positions Nucleosomes Locally. *Journal of Molecular Biology*, *391*(1), 12-25.
- Peck, L., & Wang, J. (1981). Sequence dependence of the helical repeat of DNA in solution. *Nature*, *292*(5821), 375-378.
- Pederson, D., Venkatesan, M., Thomat, F., & Simpson, R. (1986). *Isolation of an episomal yeast gene and replication origin as chromatin (chromatin structure/linking number/electron microscopy/transcription/Saccharomyces cerevisiae)*.

- Pietrasanta, L., Thrower, D., Hsieh, W., Rao, S., Stemmann, O., Lechner, J., Carbon, J., & Hansma, H. (1999). Probing the *Saccharomyces cerevisiae* centromeric DNA (CEN DNA)-Binding factor3 (CBF3) kinetochore complex by using atomic force microscopy. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(7), 3757-3762.
- Prunell, A. (1998). A topological approach to nucleosome structure and dynamics: The linking number paradox and other issues. *Biophysical Journal*, *74*(5), 2531-2544.
- Prunell, A., Kornberg, R., Lutter, L., Klug, A., Levitt, M., & Crick, F. (1979). Periodicity of deoxyribonuclease I digestion of chromatin. *Science*, *204*(4395), 855.
- Purvis, A., & Singleton, M. (2008). Insights into kinetochore-DNA interactions from the structure of Cep3Δ. *EMBO Reports*, *9*(1), 56-62.
- Radman-Livaja, M., & Rando, O. (2010). Nucleosome positioning: How is it established, and why does it matter? *Developmental Biology*, *339*(2), 258-266.
- Rando, O., & Ahmad, K. (2007). Rules and regulation in the primary structure of chromatin. *Current opinion in cell biology*, *19*(3), 250-256.
- Rando, O., & Winston, F. (2012). Chromatin and transcription in yeast. *Genetics*, *190*(2), 351-387.
- Ricci, M., Manzo, C., García-Parajo, M., Lakadamyali, M., & Cosma, M. (2015). Chromatin fibers are formed by heterogeneous groups of nucleosomes in vivo. *Cell*, *160*(6), 1145-1158.
- Richmond, T., Finch, J., Rushton, B., Rhodes, D., & Klug, A. (1984). Structure of the nucleosome core particle at 7 resolution. *Nature*, *311*(5986), 532-537.
- Robinson, P., Fairall, L., Huynh, V., & Rhodes, D. (2006). EM measurements define the dimensions of the "30-nm" chromatin fiber: Evidence for a compact, interdigitated structure. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(17), 6506-6511.
- Roca, J., & Wang, J. (1992). The capture of a DNA double helix by an ATP-dependent protein clamp: A key step in DNA transport by type II DNA topoisomerases. *Cell*, *71*(5), 833-840.
- Roca, J., & Wang, J. (1994). DNA transport by a type II DNA topoisomerase: Evidence in favor of a two-gate mechanism. *Cell*, *77*(4), 609-616.
- Roca, J., Berger, J., Harrison, S., & Wang, J. (1996). DNA transport by a type II topoisomerase: Direct evidence for a two-gate mechanism. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(9), 4057-4062.
- Rodriguez, J., McKnight, J., & Tsukiyama, T. (2014). Genome-wide analysis of nucleosome positions, occupancy, and accessibility in yeast: Nucleosome mapping, high-resolution histone ChIP, and NCAM. *Current Protocols in Molecular Biology*, *2014*, 21.28.1-21.28.16.

- Russell, I., Grancell, A., & Sorger, P. (1999). The unstable F-box protein p58-Ctf13 forms the structural core of the CBF3 kinetochore complex. *Journal of Cell Biology*, 145(5), 933-950.
- Saragosti, S., Moyne, G., & Yaniv, M. (1980). Absence of nucleosomes in a fraction of SV40 chromatin between the origin of replication and the region coding for the late leader RNA. *Cell*, 20(1), 65-73.
- Segal, E., Fondufe-Mittendorf, Y., Chen, L., Thåström, A., Field, Y., Moore, I., Wang, J., & Widom, J. (2006). A genomic code for nucleosome positioning. *Nature*, 442(7104), 772-778.
- Segal, E., & Widom, J. (2009). What controls nucleosome positions? *Trends in genetics : TIG*, 25(8), 335-43.
- Segura, J., Joshi, R., Díaz-Ingelmo, O., Valdés, A., Dyson, S., Martínez-García, B., & Roca, J. (2018). Intracellular nucleosomes constrain a DNA linking number difference of -1.26 that reconciles the Lk paradox. *Nature Communications*, 9(1), 3989.
- Sekinger, E., Moqtaderi, Z., & Struhl, K. (2005). Intrinsic histone-DNA interactions and low nucleosome density are important for preferential accessibility of promoter regions in yeast. *Molecular Cell*, 18(6), 735-748.
- Shishido, K., Komiyama, N., & Ikawa, S. (1987). Increased production of a knotted form of plasmid pBR322 DNA in Escherichia coli DNA topoisomerase mutants. *Journal of Molecular Biology*, 195(1), 215-218.
- Shure, M., & Vinograd, J. (1976). The number of superhelical turns in native Virion SV40 DNA and Minicol DNA determined by the band counting method. *Cell*, 8(2), 215-226.
- Simpson, R. (1978). Structure of the chromatosome, a chromatin particle containing 160 base pairs of DNA and all the histones. *Biochemistry*, 17(25), 5524-5531.
- Simpson, R., Thoma, F., & Brubaker, J. (1985). Chromatin reconstituted from tandemly repeated cloned DNA fragments and core histones: A model system for study of higher order structure. *Cell*, 42(3), 799-808.
- Sogo, J., Stahl, H., Koller, T., & Knippers, R. (1986). Structure of replicating simian virus 40 minichromosomes: The replication fork, core histone segregation and terminal structures. *Journal of Molecular Biology*, 189(1), 189-204.
- Talbert, P., & Henikoff, S. (2010). Histone variants ancient wrap artists of the epigenome. *Nature Reviews, Molecular Cell Biology*, 11(4), 264-275.
- Tanaka, S., Livingstone-Zatchej, M., & Thoma, F. (1996). Chromatin structure of the yeast URA3 gene at high resolution provides insight into structure and positioning of nucleosomes in the chromosomal context. *Journal of Molecular Biology*, 257(5), 919-934.

- Tomschik, M., Zheng, H., Van Holde, K., Zlatanova, J., & Leuba, S. (2005). Fast long-range, reversible conformational fluctuations in nucleosomes revealed by single-pair fluorescence resonance energy transfer. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(9), 3278-3283.
- Travers, A., & Klug, A. (1987). The bending of DNA in nucleosomes and its wider implications. *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences*, *317*(1187), 537-561.
- Turner, B. (2007). Defining an epigenetic code. *Nature Cell Biology*, *9*(1), 2-6.
- Valdés, A., Segura, J., Dyson, S., Martínez-García, B., & Roca, J. (2018). DNA knots occur in intracellular chromatin. *Nucleic Acids Research*, *46*(2), 650-660.
- Valouev, A., Ichikawa, J., Tonthat, T., Stuart, J., Ranade, S., Peckham, H., Zeng, K., Malek, J., Costa, G., Mckernan, K., Sidow, A., Fire, A., & Johnson, S. (2008). A high-resolution, nucleosome position map of *C. elegans* reveals a lack of universal sequence-dictated positioning. *Genome Research*, *18*(7), 1051-1063.
- Van Holde, K., & Zlatanova, J. (2006). Scanning chromatin. A new paradigm? *Journal of Biological Chemistry*, *281*(18), 12197-12200.
- Vinograd, J., Lebowitz, J., Radloff, R., Watson, R., & Laipis, P. (1965). THE TWISTED CIRCULAR FORM OF POLYOMA VIRAL DNA*. *Proceedings of the National Academy of Sciences of the United States*, *53*(5), 1104-1111.
- Vlijm, R., Lee, M., Lipfert, J., Lusser, A., Dekker, C., & Dekker, N. (2015). Nucleosome assembly dynamics involve spontaneous fluctuations in the handedness of tetrasomes. *Cell Reports*, *10*(2), 216-225.
- Vologodskii, A., & Cozzarelli, N. (1994). Conformational and Thermodynamic properties of supercoiled DNA. *Annual Review of Biophysics and Biomolecular Structure*, *23*, 609-643.
- Vos, S., Tretter, E., Schmidt, B., & Berger, J. (2011). All tangled up: How cells direct, manage and exploit topoisomerase function. *Nature Reviews. Molecular Cell Biology*, *12*(12), 827-841.
- Wallis, J., Chrebet, G., Brodsky, G., Rolfe, M., & Rothstein, R. (1989). A hyper-recombination mutation in *S. cerevisiae* identifies a novel eukaryotic topoisomerase. *Cell*, *58*(2), 409-419.
- Wang, J. (1971). Interaction between DNA and an *Escherichia coli* protein ω . *Journal of Molecular Biology*, *55*(3), 523-533.
- Wang, J. (1979). Helical repeat of DNA in solution. *Proceedings of the National Academy of Sciences of the United States of America*, *76*(1), 200-203.
- Wang, J. (1996). DNA Topoisomerases. *Annual Review of Biochemistry*, *65*(1), 635-692.

- Washietl, S., Machné, R., & Goldman, N. (2008). Evolutionary footprints of nucleosome positions in yeast. *Trends in genetics*, 24(12), 583-587.
- Watson, J., & Crick, F. (1953). Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. *Nature*, 171, 737-738.
- White, J. (1969). Self-Linking and the Gauss Integral in Higher Dimensions. *American Journal of Mathematics*, 91(3), 693.
- White, J., & Bauer, W. (1989). The helical repeat of nucleosome-wrapped DNA. *Cell*, 56(1), 9-10.
- Wieland, G., Orthaus, S., Ohndorf, S., Diekmann, S., & Hemmerich, P. (2004). Functional Complementation of Human Centromere Protein A (CENP-A) by Cse4p from *Saccharomyces cerevisiae*. *Molecular and Cellular Biology*, 24(15), 6620-6630.
- Wisniewski, J., Hajj, B., Chen, J., Mizuguchi, G., Xiao, H., Wei, D., . . . Wu, C. (2014). Imaging the fate of histone Cse4 reveals de novo replacement in S phase and subsequent stable residence at centromeres. *eLife*, 2014(3).
- Worcel, A., Strogatz, S., & Riley, D. (1981). Structure of chromatin and the linking number of DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 78(3), 1461-1465.
- Xiao, H., Mizuguchi, G., Wisniewski, J., Huang, Y., Wei, D., & Wu, C. (2011). Nonhistone Scm3 Binds to AT-Rich DNA to Organize Atypical Centromeric Nucleosome of Budding Yeast. *Molecular Cell*, 43(3), 369-380.
- Yarragudi, A., Miyake, T., Li, R., & Morse, R. (2004). Comparison of ABF1 and RAP1 in Chromatin Opening and Transactivator Potentiation in the Budding Yeast *Saccharomyces cerevisiae*. *Molecular and Cellular Biology*, 24(20), 9152-9164.
- Yuan, G., Liu, Y., Dion, M., Slack, M., Wu, L., Altschuler, S., & Rando, O. (2005). Molecular biology: Genome-scale identification of nucleosome positions in *S. cerevisiae*. *Science*, 309(5734), 626-630.
- Zentner, G., & Henikoff, S. (2013). Regulation of nucleosome dynamics by histone modifications, *Nature Structural and Molecular Biology*, 20(3), 259-266.
- Zhang, W., Lukoyanova, N., Miah, S., Lucas, J., & Vaughan, C.K. (2019). Insights into centromere DNA bending revealed by the Cryo-EM Structure of the core centromere binding factor 3 with Ndc10. *Cell Reports*, 26(4), 1070.
- Zivanovic, Y., Goulet, I., Revet, B., Le Bret, M., & Prunell, A. (1988). Chromatin reconstitution on small DNA rings. II. DNA supercoiling on the nucleosome. *Journal of Molecular Biology*, 200(2), 267-285.

Zlatanova, J., Bishop, T., Victor, J.-M., Jackson, V., & van Holde, K. (2009). The Nucleosome Family: Dynamic and Growing. *Structure*, 17(2), 160-171.

Zubay, G., & Doty, P. (1959). The isolation and properties of deoxyribonucleoprotein particles containing single nucleic acid molecules. *Journal of Molecular Biology*, 1(1), 1-20.

