CENP-B-mediated DNA loops regulate activity and stability of human centromeres

Chardon, Florian; Japaridze, Aleksandre; Witt, Hannes; Velikovsky, Leonid; Chakraborty, Camellia; Wilhelm, Therese; Dumont, Marie; Yang, Wayne; Dekker, Cees; More Authors

DOI
10.1016/j.molcel.2022.02.032

Publication date
2022

Document Version
Final published version

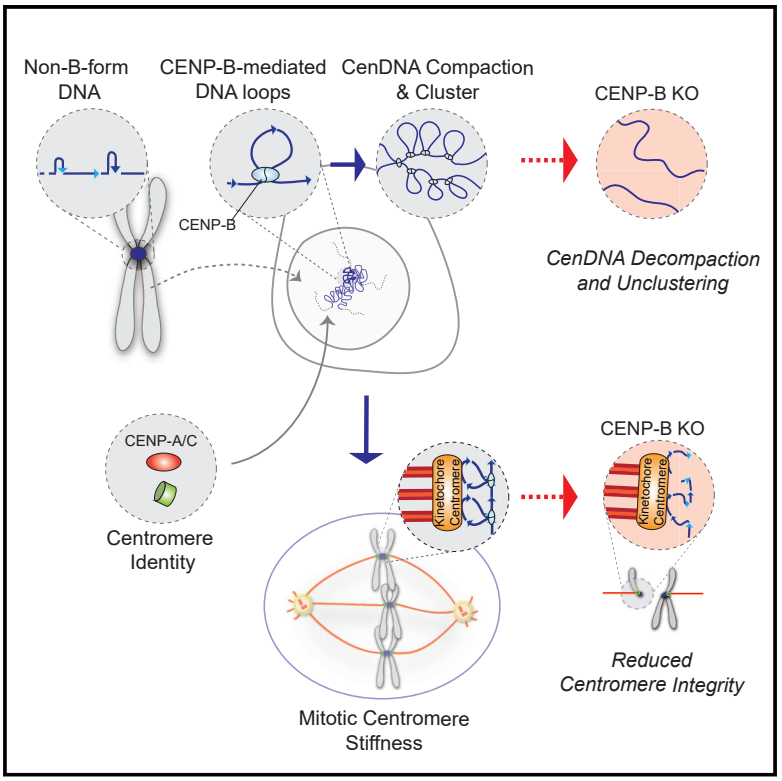
Published in
Molecular Cell

Molecular Cell

Article

CENP-B-mediated DNA loops regulate activity and stability of human centromeres

Graphical abstract



Highlights

- Centromeric DNA self-organizes into non-B-form DNA structures like hairpins
- CENP-B reshapes centromeres by forming sub-micron-sized DNA loops between DNA repeats
- CENP-B-mediated DNA loops compact centromeric chromatin and promote clustering
- Maintenance of CENP-B-induced architecture favors centromere identity and integrity

Authors

Florian Chardon, Aleksandre Japaridze, Hannes Witt, ..., Gijs Wuite, Cees Dekker, Daniele Fachinetti

Correspondence

daniele.fachinetti@curie.fr

In brief

Chardon et al. discovered, using various single-molecule techniques, that centromeric DNA self-organizes into non-B-form DNA including hairpins and that CENP-B promotes DNA loop formation within centromeric repeats. This CENP-B-mediated architecture is important for the organization of centromeric chromatin and favors centromere maintenance and integrity in human cells.

Chardon et al., 2022, Molecular Cell *82*, 1751–1767
May 5, 2022
https://doi.org/10.1016/j.molcel.2022.02.032

CellPress

Molecular Cell

CellPress

Article

CENP-B-mediated DNA loops regulate activity and stability of human centromeres

Florian Chardon,[1,4] Aleksandre Japaridze,[2,4] Hannes Witt,[3] Leonid Velikovsky,[1] Camellia Chakraborty,[1] Therese Wilhelm,[1] Marie Dumont,[1] Wayne Yang,[2] Carlos Kikuti,[1] Stephane Gangnard,[1] Anne-Sophie Mace,[1] Gijs Wuite,[3] Cees Dekker,[2] and Daniele Fachinetti[1,5,*]

[1]Institut Curie, PSL Research University, CNRS, UMR 144, 26 rue d'Ulm, 75005 Paris, France

[2]Department of Bionanoscience, Kavli Institute of Nanoscience Delft, Delft University of Technology, Van der Maasweg 9, 2629 HZ Delft, the Netherlands

[3]Department of Physics and Astronomy, LaserLaB Amsterdam, Vrije Universiteit Amsterdam, De Boelelaan 1081, 1081 HV Amsterdam, the Netherlands

[4]These authors contributed equally

[5]Lead contact

*Correspondence: daniele.fachinetti@curie.fr

https://doi.org/10.1016/j.molcel.2022.02.032

SUMMARY

Chromosome inheritance depends on centromeres, epigenetically specified regions of chromosomes. While conventional human centromeres are known to be built of long tandem DNA repeats, much of their architecture remains unknown. Using single-molecule techniques such as AFM, nanopores, and optical tweezers, we find that human centromeric DNA exhibits complex DNA folds such as local hairpins. Upon binding to a specific sequence within centromeric regions, the DNA-binding protein CENP-B compacts centromeres by forming pronounced DNA loops between the repeats, which favor inter-chromosomal centromere compaction and clustering. This DNA-loop-mediated organization of centromeric chromatin participates in maintaining centromere position and integrity upon microtubule pulling during mitosis. Our findings emphasize the importance of DNA topology in centromeric regulation and stability.

INTRODUCTION

Centromeres are essential regions of chromosomes acting as platforms for kinetochore assembly and microtubule (MT) attachment to guide proper genome repartitioning during both mitosis and meiosis (Fukagawa and Earnshaw, 2014; Mellone and Fachinetti, 2021). Despite being epigenetically regulated, most centromeres across multiple species are enriched with particular types of AT-rich DNA sequences, raising questions on the functionality of such repeats (Dumont and Fachinetti, 2017; Talbert and Henikoff, 2020).

In human cells, centromeres are built with a unique type of DNA tandem repeats, called α-satellites (α-sat) (Wu and Manuelidis, 1980), organized in a head-to-tail configuration of single AT-nucleotide-rich 171-bp monomers with high sequence similarity (Altemose et al., 2021). These monomers form larger repetitive units named higher order repeats (HORs) that span up to 5 Mb and represent ~3% of the total genome (Logsdon et al., 2021; Miga et al., 2020). Within roughly one in every other α-sat repeat there is a unique, and highly conserved, 17-bp DNA motif named centromeric protein B (CENP-B) box, bound by CENP-B (Earnshaw et al., 1987; Gamba and Fachinetti, 2020; Masumoto et al., 1989; Muro et al., 1992; Yoda et al., 1992). CENP-B originates from the domestication of pogo-like transposases (Casola et al., 2008), and it is the only known centromeric DNA-sequence-specific binding protein that associates with the α-sat of all centromeres (except the Y chromosome; Earnshaw et al., 1987).

DNA repeats are known to self-assemble into unusual structures through base-pairing (Mirkin, 2007). Given its repetitive nature, AT-rich centromeric DNA is likely no exception. Accordingly, *in silico* predictions and NMR spectroscopy studies on satellite DNA from various species have suggested the presence of short DNA dyad symmetries (Kasinathan and Henikoff, 2018) that can favor the formation of several types of non-B-form DNA, such as hairpins (Gallego et al., 1997; Jonstrup et al., 2008), Z-DNA (Li et al., 2009), and intercalated motifs (i-motifs) (Garavís et al., 2015; Nonin-Lecomte and Leroy, 2001). Formation of secondary structures were not predicted in centromeres containing CENP-B boxes, suggesting that CENP-B itself could be implicated in promoting secondary structures once bound to centromeric DNA (Kasinathan and Henikoff, 2018). CENP-B was shown to have DNA-bending properties leading to a translational positioning of the nucleosome (Tanaka et al., 2001, 2005a; Yoda et al., 1992). Furthermore, the distribution of CENP-B boxes suggests that CENP-B influences the positioning of centromeric

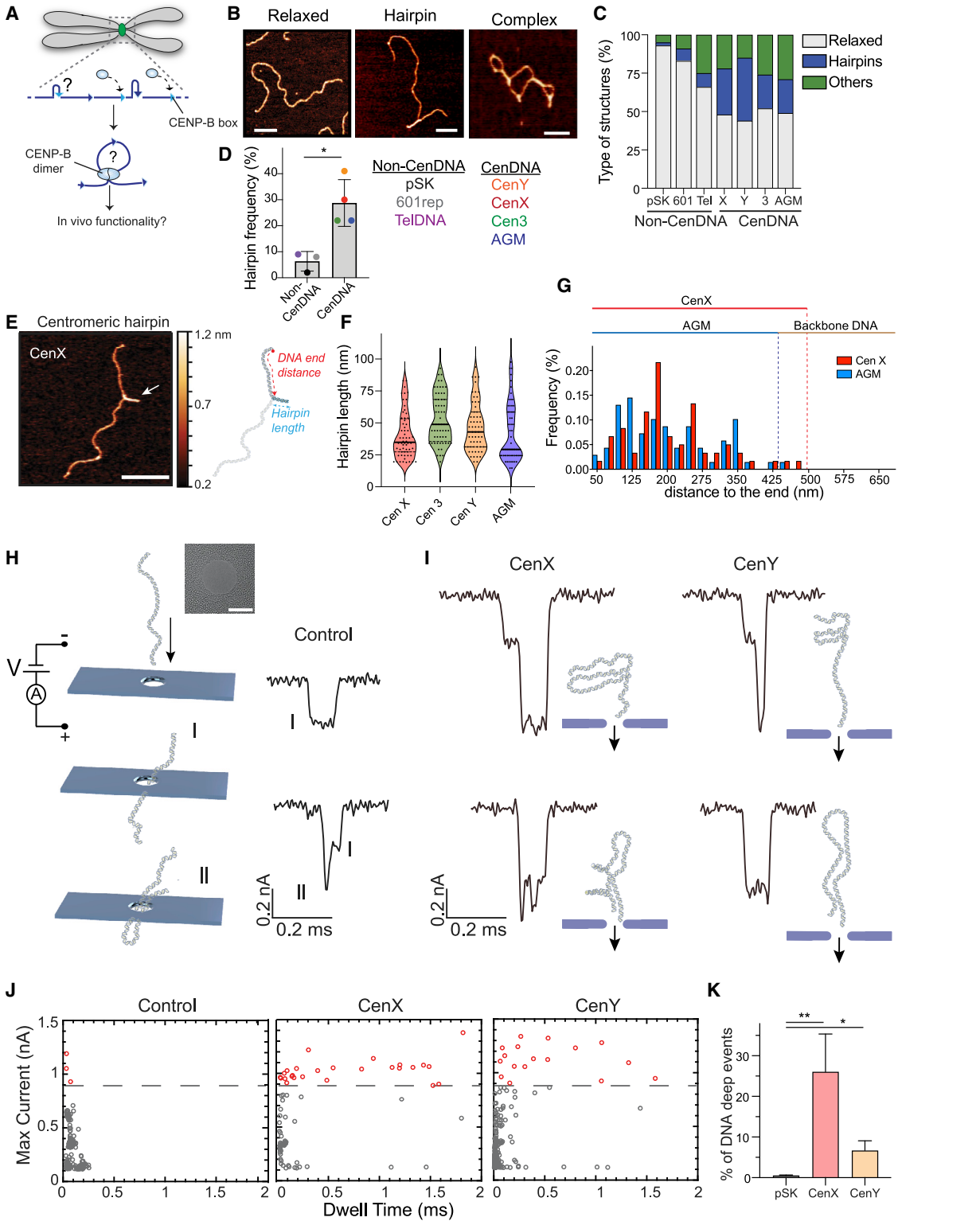


Figure 1. Centromeric DNA has higher propensity to form secondary structures

(A) Schematic of the scope of the manuscript: investigate the presence of secondary structures (depicted as bended arrows) within centromeric DNA (straight arrows) and assess if/how CENP-B (light blue circles) reorganizes centromeric DNA once it is bound to CENP-B boxes (small blue triangles) and, if so, the functionality *in vivo*.

(legend continued on next page)

nucleosomes (Hasson et al., 2013; Nechemia-Arbely et al., 2017) and overall chromatin fiber architecture (Tawaramoto et al., 2003; Yoda et al., 1998).

Whether human centromeres fold into unique secondary DNA structures and how CENP-B binding remodels the overall repetitive centromeric architecture are still key unsolved questions (Figure 1A). Secondary structures have been implicated in maintenance of centromere identity (Kasinathan and Henikoff, 2018; Mellone and Fachinetti, 2021) and stability (Black and Giunta, 2018; Bloom and Costanzo, 2017), but clear experimental demonstrations of their existence and roles in human cells are missing. Here, we use a wide range of single-molecule techniques to show that human centromeric DNA self-organizes into non-B-form secondary DNA structures. Binding of CENP-B to CENP-B boxes is then capable of reshaping centromeres by forming sub-micron sized DNA loops between repeats that compact and cluster centromeric chromatin. Further, we show that CENP-B-mediated DNA looping is a highly important feature to preserve centromere positioning and DNA integrity in living cells.

RESULTS

Human centromeric DNA forms secondary structures

To directly interrogate the existence of non-B-form DNA in human centromeres, we generated centromeric DNA (17 α-sat monomers, $\sim$3 kb) derived from the two human sex chromosomes (hereafter referred to as CenY and CenX) and chromosome 3 (Cen3) (Figure S1A). All these sequences are known to be bound by CENP-A *in vivo* (Dumont et al., 2020; Skene and Henikoff, 2015). To directly visualize whether centromeric DNA *per se* was prone to form non-canonical DNA structures at nanometer resolution, we first used atomic force microscopy (AFM) (Binnig et al., 1986; Japaridze et al., 2016). The measured contour length (1,000 nm) and the apparent DNA persistence length (50 nm)—a measure describing the stiffness of the DNA molecule (Rubinstein and Colby, 2003; Figure S1B; Table S1)—of linear centromeric DNA deposited on mica surface were in good agreement with previously reported values for B-form DNA (Japaridze et al., 2016, 2017; Baumann et al., 1997).

Surprisingly, centromeric DNA was found to preferentially form local double-stranded (ds) hairpins and complex secondary structures rather than merely unstructured relaxed B-form DNA (>50% of all structures; Figures 1B–1D). These pronounced features were clearly due to the DNA sequence and not caused by sample preparation, as non-repetitive DNA (pSK 3-kb linearized plasmid lacking centromeric sequences) or repetitive but not-centromeric DNA (12 tandem repeats of 208-bp Widom 601 sequence or telomeric DNA) only displayed relaxed conformations (>90% of all structures in pSK) or other structures with only few hairpins (<10% of all structures) (Figures 1C, 1D, S1A, and S1B). In addition, hairpin accumulation in centromeric DNA was not due to Mg^{2+} ions in the buffer (known to affect secondary structure), as changing the Mg^{2+} concentration did not affect their formation (Figure S1C). The formation of hairpins and more complex secondary structures were also observed on DNA derived from CenY and from the primate *Chlorocebus sabeus* (AGM, African green monkey) (Figures 1C, 1D, S1A, and S1B), both lacking CENP-B boxes and predicted to form secondary structures *in silico* (Kasinathan and Henikoff, 2018). This shows that these non-canonical DNA structures formed independently of the presence of CENP-B boxes and are a common feature of centromeric DNA. Hairpin length was similar across the different centromeric DNAs (CenX: 41 ± 18 nm, Cen3: 53 ± 19 nm, CenY: 47 ± 19 nm, AGM: 42 ± 23 nm; mean ± SD), and hairpins were distributed unevenly along the whole range of centromeric DNA but not outside of it (Figures 1E–1G). This indicates that hairpin structures can form within the same α-sat monomer or involve two consecutive ones (median of 106–147 bp; 1 bp = $\sim$0.33 nm; Ghosh and Bansal, 2003) and that they could form at several sites along the DNA molecules.

To confirm these AFM results, we investigated the formation of secondary structures of centromeric DNA molecules in solution by using solid-state nanopores (Japaridze et al., 2021; Plesa et al., 2016; Kumar Sharma et al., 2019). The method is based on electrophoretic voltage-driven translocation of DNA molecules across a nanometer-sized pore (7-nm diameter in our case) (Figure 1H). The translocation of the DNA across the pore temporarily disrupts the flow of ions which leads to a

(B) Representative zoomed AFM images of relaxed, hairpin, and complex DNA structures. Scale bars, 100 nm.
(C) Bar graph depicting the relative frequency of the type of structures shown in (B). Control SK DNA (N = 107), CenX (N = 139), CenY (N = 80), Cen3 (N = 120), 601 (N = 318), AGM (N = 207), and TelDNA (N = 360). The category "others" includes DNA loops and complex structures. See Table S1 for details.
(D) Bar graph shows hairpin frequency on the different DNA fragments. Unpaired t test *p = 0.0105. See Table S1 for details. Error bars show the SD.
(E) AFM image of a hairpin DNA formed on CenX DNA with corresponding schematic depiction of the hairpin distance from the end and the hairpin length. Scale bars, 200 nm.
(F) Violin plot showing hairpin length in the indicated DNA. Bars represent median, dotted bars quartiles. See Table S1 for details.
(G) Frequency of CenX (red) or AGM (blue) hairpin distances from either one of the two DNA ends (N = 55 and 70, respectively). Red dotted line marks length of CenX, while blue and brown dotted lines mark AGM centromeric DNA or backbone, non-centromeric DNA. See Table S1 for details.
(H) Schematic depiction of the nanopore setup and various types of control DNA translocation events. Type I events indicate translocation of an unfolded dsDNA; type II events indicate a double dsDNA fold. Inset shows the TEM image of the 7-nm pore used in the measurements. Scale bars, 5 nm.
(I) Two examples of CenX (left) and CenY (right) current traces along with schematic of how the DNA may be folded. Black arrows indicate the DNA translocation direction.
(J) Maximum current blockade distribution for control 3-kb DNA (N = 496) (left), CenX DNA (N = 87) (middle), and CenY DNA (N = 279) (right). Horizontal dotted lines indicate the event threshold set at 3 times the average current blockade of a single DNA helix (3xIo). Red circles indicate the DNA events that are above the threshold level.
(K) Histogram showing the percentage of deep events from (J) that are larger than the threshold level. Deep events are seen only on both CenX and CenY, but not on control DNA. Error bars show the SD. Unpaired t test * p = 0.0112; ** p = 0.0089.

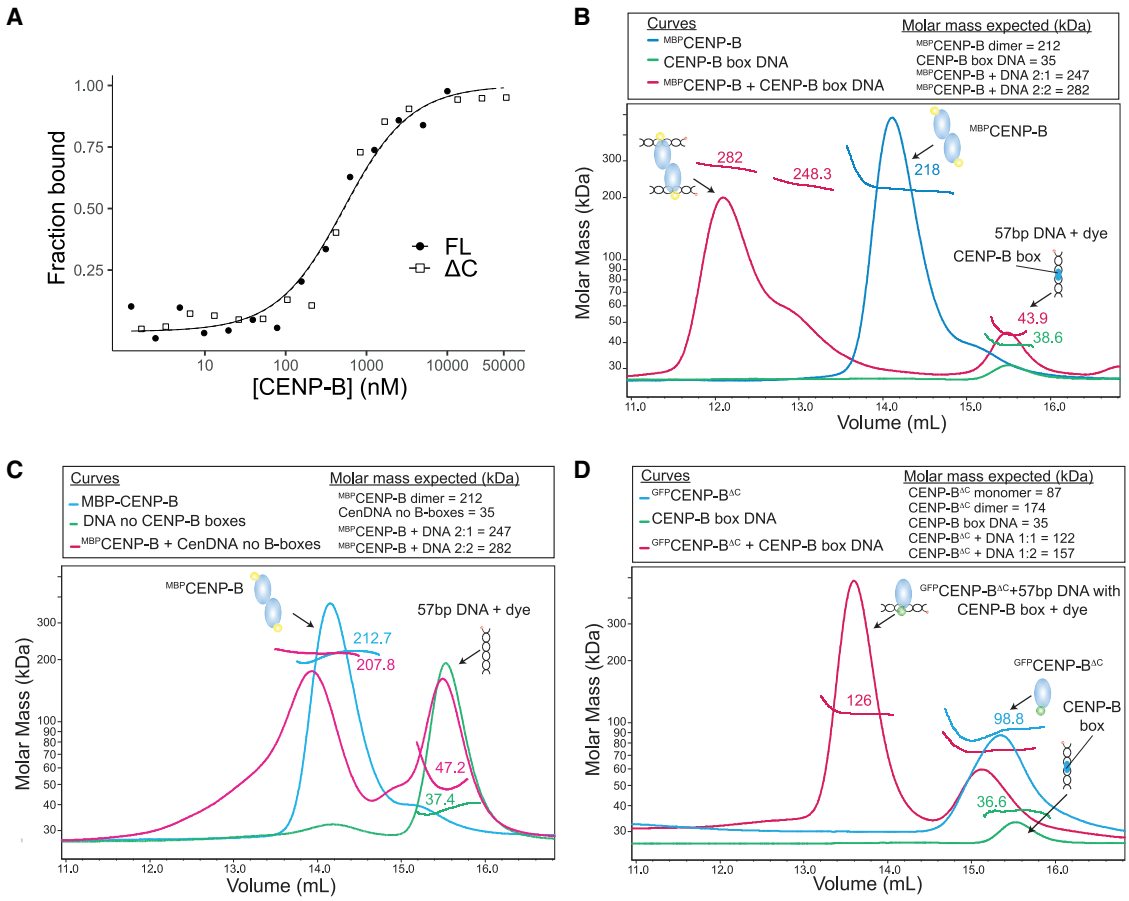


Figure 2. CENP-B proteins form a tetrameric complex with centromeric DNA

(A) Representative microscale thermophoresis (MST) experiment to measure the affinity between FL-CENP-B (dots) or CENP-B$^{\Delta C}$ (squares) and a 57-bp DNA fragment containing 1 CENP-B box. Lines represent binding curves fitted using Kd-binding model. See Table S1 for details.

(B–D) SEC-MALS analyses identify the stoichiometry of the indicated complexes. Blue lines: light-scattering profiles (3 angles combined) and molecular masses of CENP-B either full-length (B) or delta dimerization domain (ΔC) (D); green lines: light-scattering and mass profiles of the 57-bp DNA containing (B and D) or not (C) the CENP-B box sequence; red lines: profiles obtained when CENP-B and the DNA fragment are mixed before injection. Here, we can observe: CENP-B dimers in complex with two CenDNA molecules (B), CENP-B dimers that did not interact with DNA (C), and a single molecule of CENP-B that interacts with only one molecule of CenDNA (D). The expected molecular masses (in kDa) are indicated. The measured kDa were calculated with Astra 6.1.7 (Wyatt).

drop in the measured current. Based on the magnitude of these current blockades, it is possible to estimate the folding of DNA molecules since multiple DNA helices that traverse the nanopore in parallel yield a deeper blockade than a single helix (Plesa et al., 2016; Kumar Sharma et al., 2019). As expected, the control 3-kb DNA showed translocation events of so-called type I and type II (Plesa and Dekker, 2015; Storm et al., 2005), which are the result of a single dsDNA or a single folded DNA that is translocating through a pore (Figures 1H, 1J, and S1D). In contrast to the control, both CenX and CenY showed additional, much deeper, and more heterogeneous blockade events, indicating complex DNA folding in solution (Figures 1I, 1J, S1D, and S1E). By setting the threshold 3 times the average current blockade of a single DNA helix, we observed that only ~1% of all molecules displayed deep translocations in control DNA, while this percentage was significantly higher in centromeric DNA (27.6% and 6.1% for CenX and CenY, respectively) (Figures 1J and 1K).

Taken together, these results indicate that human centromeric DNA is folded into non-B-form DNA as hairpins and more complex structures.

Homodimers of CENP-B bind to CENP-B boxes forming a tetrameric complex

Next, we wanted to examine the capability of CENP-B to bind and reshape centromeric DNA. We first expressed and purified full-length (FL) CENP-B and its dimerization mutant (ΔC) (Figures S2A and S2B). CENP-B$^{\Delta C}$ has the capacity to bind and bend DNA (Tanaka et al., 2001), but it cannot dimerize (and hence likely cannot form any secondary structure). Both FL-CENP-B and CENP-B$^{\Delta C}$ bind to a short DNA sequence containing one CENP-B box motif with the same apparent K_D (0.5 μM) as determined by microscale thermophoresis (MST) (Figure 2A). To determine the absolute molar masses and the stoichiometry of this interaction, we performed size-exclusion chromatography

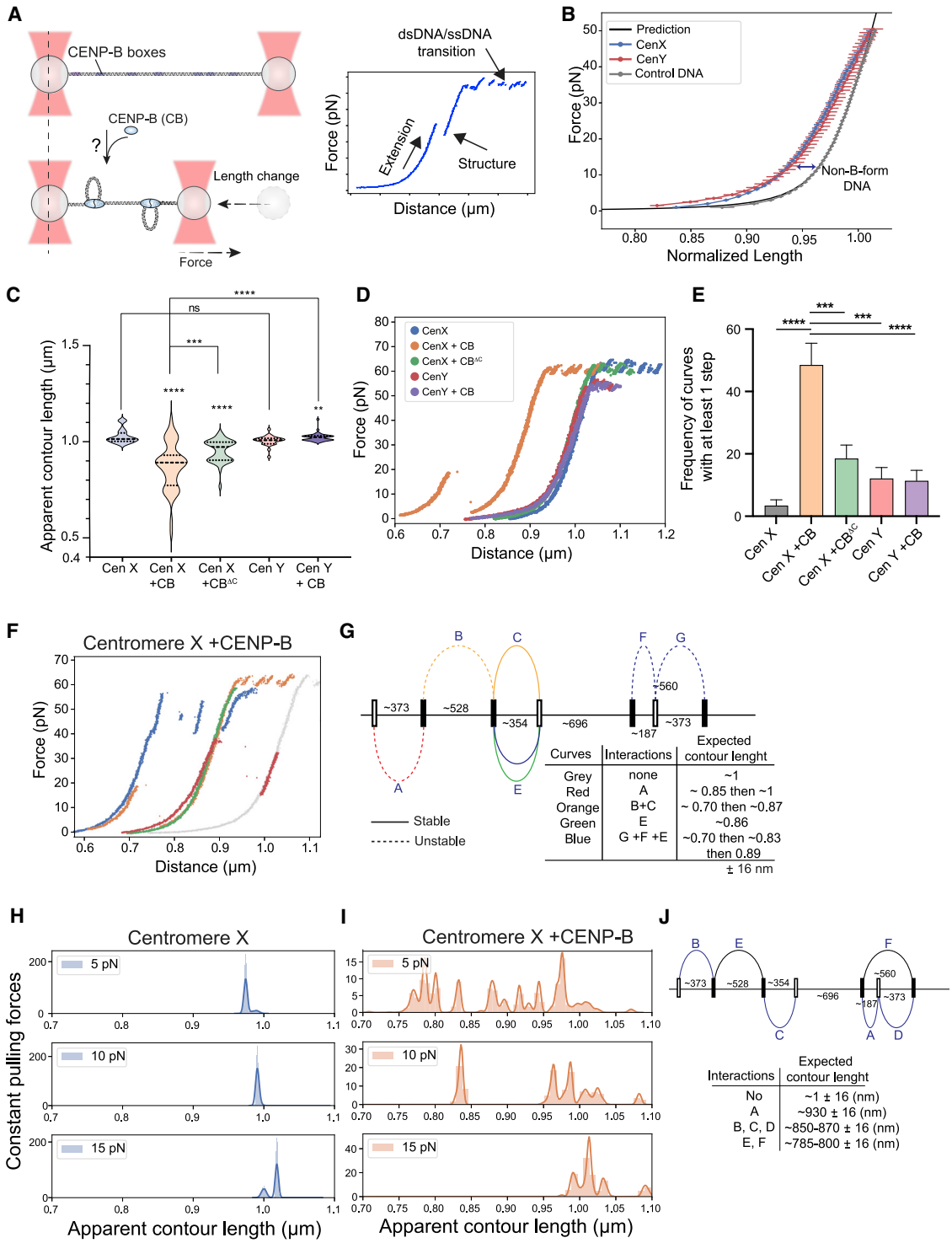


Figure 3. High-resolution optical tweezers detect DNA changes due to homodimer CENP-B binding to CENP-B boxes

(A) Schematic of the experiment. DNA is clamped between two beads held by optical traps. Loop formation induced by CENP-B is detected as a change of the DNA end-to-end length. An example of force-distance curve obtained after pulling a structured DNA is also shown.

(B) Averaged and normalized stretch curves for CenX, CenY, and PKYB1 (control) DNA. The predicted behavior for dsDNA (black line) is also shown. Error bars show the SEM.

(C) Violin plots show the distribution of the determined contour length in the indicated conditions. Mann-Whitney U test, ns: 0.05 < p, **: 0.001 < p < 0.01, ***: 0.0001 < p < 0.001, ****: p < 0.0001. See Table S1 for details.

(legend continued on next page)

multi-angle light-scattering (SEC-MALS) analysis (Figure 2B). FL-CENP-B protein (blue curve) formed a homodimer as revealed by its molecular weight (MW of 218 kDa), while the molar mass of DNA alone (green curve) is close to the expected value (38.6 kDa) (Figure 2B). Pre-incubation of FL-CENP-B with DNA (red curve) revealed a main peak at 282 kDa corresponding to a dimer of CENP-B in complex with two molecules of DNA. We observed two additional peaks, corresponding to the excess of DNA (43.9 kDa), and a shoulder on the main peak (248.3 kDa) that could correspond to a transition state with a dimer of CENP-B complexed with one DNA molecule (Figure 2B). This formation of a CENP-B/DNA tetrameric complex was specific to the presence of the CENP-B box and dependent on the dimerization domain of CENP-B (Figures 2C and 2D). In summary, these data show that a CENP-B dimer specifically interacts with two different molecules of DNA containing CENP-B boxes through its dimerization and DNA-binding domain.

We then assessed if and how CENP-B shapes the topology of the centromeric DNA. To test the binding properties of CENP-B to DNA, we used CenX and CenY, the latter serving as a negative control since it lacks CENP-B boxes and CENP-B cannot bind to it. Within the 17 α-sat monomers, CenX contained 7 potential CENP-B-binding sites (4 canonical CENP-B boxes consensus sequences of NTTCGNNNNANNCGGGN and 3 putative ones NTTCNNNNNANNCGGGN) (Figure S2C). The electrophoretic mobility shift assay (EMSA) on CenX revealed that FL-CENP-B and CENP-B$^{\Delta C}$ were both capable of binding to DNA (Figure S2D) in a concentration-dependent manner (Figure S2E). As previously reported (Yoda et al., 1998), CENP-B binds several CENP-B boxes and promotes DNA complexes as judged by the presence of a higher MW smear that disappeared with the removal of the CENP-B dimerization domain (Figure S2D). Similar results were obtained on CenX derivatives (CenX-300, -500 and -700) that carry three, four, or five monomers of α-sat DNA embedded in non-centromeric DNA with varying distances between the two CENP-B boxes (Figures S2F–S2H). Here, we could also detect a band with about twice the MW of a single DNA molecule, which likely represents two molecules of DNA brought together by CENP-B via its dimerization domain (Figures S2G and S2H).

Homodimers of CENP-B reshape centromeric DNA

To assess the architecture of centromeric DNA at a single-molecule level, we employed high-resolution optical tweezers that enable DNA manipulation with sub-nanometer spatial and piconewton (pN) force accuracy to obtain force/distance curves. A single dsDNA of CenX or CenY was attached at both ends to beads (Figure 3A). These beads were then manipulated by two optical traps while the force response and the length of the DNA were measured (Figures 3A and S3A). When CenX or CenY were stretched at a constant velocity, we observed the typical force-distance curve for DNA with a non-linear force increase expected for a worm-like chain (WLC) followed by a typical force plateau at ~65 pN indicating overstretching of the dsDNA helix (Figure 3B; Smith et al., 1996). However, when comparing their stretch response to the prediction (WLC with a persistence length of ~50 nm, as measured by AFM; Table S1), it became evident that centromeric DNA, but not non-centromeric DNA (PKYB1 DNA), got softer, which led to a force-dependent shortening with a length decrease of up to ~30 nm at ~10 pN (Figures 3B and S3B). Quantitatively, when fitting an extensible WLC to the initial segment (up to 30 pN) of the force-distance curve (Odijk, 1995), this softening is the result of a strong reduction of the apparent persistence length (26 ± 12 nm for CenX, 17 ± 7 nm for CenY, 43 ± 11 nm for PKYB1, mean ± SD) (Figures S3C and S3D). Since the persistence length of dsDNA under stable buffer conditions is highly reproducible (Gross et al., 2011), we relate the force-dependent shortening of centromeric DNA to the occurrence of non-B-form DNA as observed in the AFM and nanopore assays.

We then tested how addition of CENP-B impacts the structure of centromeric DNA. For this we first quantified the apparent DNA contour length as obtained from the WLC fit with addition of FL-CENP-B or CENP-B$^{\Delta C}$ at non-saturating condition. The measured apparent contour length of the CenX was found to be reduced by 100–300 nm upon FL-CENP-B binding (Figures 3C, S3E, and S3F; Table S1). The distribution of the apparent contour length of CenX in the presence of CENP-B showed distinct peaks with maxima at 0.75, 0.9, and 1 μm (Figures S3E and S3F). This reduction in length was not observed for CenX alone or CenY ± FL-CENP-B that instead showed only a single peak at ~1 μm, representing the length of the bare DNA. Since CenX has multiple binding sites for CENP-B, the two peaks around 0.75 and 0.9 μm could correspond to the formation of several structures (as DNA loops) between different CENP-B sites, which were larger than what observed by AFM (as hairpin in Figure 1). Since binding of proteins to DNA could also lead to shortening of the DNA helix, which can be measured directly using optical tweezers (Heller et al., 2014), we tested changes in CenX following incubation with CENP-B$^{\Delta C}$. CENP-B$^{\Delta C}$ led indeed to partial shortening of the DNA (~100 nm) due to its binding proprieties and the presence of several CENP-B boxes; however, very few curves had a contour length shorter than 0.8 μm (Figures 3C, S3E, and S3F; Table S1). In contrast, for

(D) Representative force-distance curves obtained in the indicated conditions.
(E) Bar graph showing the frequency of the occurrence of discontinuities in the stretching curve indicating the formation of secondary structure. Square root for each sample is also indicated. χ^2 test: ***: 0.0001 < p < 0.001, ****: p < 0.0001. See Table S1 for details.
(F) Examples of 5 color coded force-distance curves of CenX + FL-CENP-B. The gray curve shows a stretch curve of CenX in the absence of CENP-B.
(G) Map of possible interactions between CENP-B boxes sites derived from the data shown in (F). Canonical B boxes are depicted in black, potential B boxes in open black. Colored lines mark interaction between potential B box sites. Distance between B boxes is indicated.
(H and I) Force-clamp experiments of CenX without (H) or with (I) CENP-B. The y axis represents the probability density function and area of the histogram is normalized to 1.
(J) Map of possible interactions between CENP-B boxes sites derived from the data shown in (I). Canonical B boxes are depicted in black, potential B boxes in open black. Blue lines mark interaction between potential B box sites. Distance between B boxes is indicated.

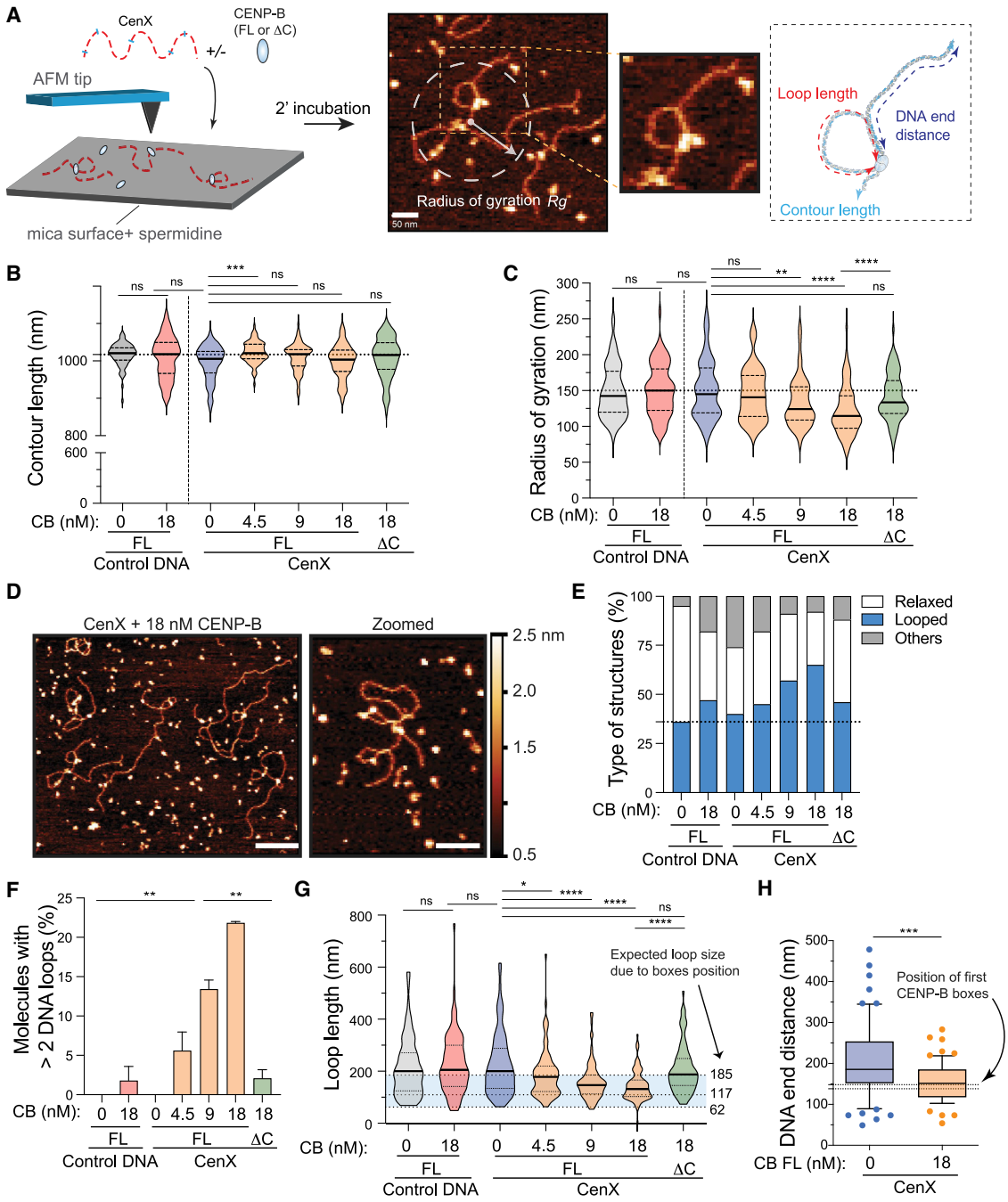


Figure 4. CENP-B promotes loop formation on centromeric DNA via its dimerization domain

(A) (Left) Schematic of the AFM experimental setting. (Center) Representative AFM image of a looped CenX bound by CENP-B. Scale bars, 50 nm. (Right) Schematics depicting the various statistical properties measured for the DNA molecules and for DNA loops: total DNA contour length (cyan dotted line), radius of gyration (white arrow), loop length (red dotted line), and distance to DNA end (blue dotted line).

(B) Violin plots showing the distribution of the control DNA or CenX DNA contour length with or without full-length (FL) or dimerization mutant (ΔC) CENP-B. Horizontal dashed line represents the mean in control DNA without CENP-B. Bars represent median, dotted bars quartiles. See Table S1 for details.

(C) Violin plots showing the distribution of the DNA radius of gyration in the indicated conditions. Horizontal dashed line marks the mean in control DNA without CENP-B. Bars represent median, dotted bars quartiles. See Table S1 for details.

(D) Representative large scale (left) and zoomed (right) AFM images of CenX DNA + CENP-B complexes. CENP-B forms multiple loops when bound to the centromeric DNA. Scale bars, 200 and 100 nm, respectively.

(E) Bar graph depicting the relative frequency of the type of structures per DNA fragment in the indicated conditions. Horizontal dashed line represents the number of looped DNA molecules in the control DNA without CENP-B. See Table S1 for details.

(legend continued on next page)

CenX + FL-CENP-B 30% of the curves fell into this category of shorter contour length. In summary, our data show that binding of CENP-B to the CENP-B boxes and its subsequent dimerization leads to a compaction of the centromeric DNA (measured as the apparent contour length), likely due to the formation of secondary structures.

During stretching of CenX DNA in the presence of FL-CENP-B, we observed steps in which the force suddenly dropped in about 50% of cases, while this was rarely observed in the other conditions (Figures 3D and 3E). Changes in the DNA length during the jumps were directly related to the position of CENP-B boxes (Figures 3F and 3G), which indicates that these steps were most likely related to the opening of CENP-B-mediated secondary structures. These structures form only between two consecutive CENP-B boxes (from ~187 to ~560 bp [~56 to ~168 nm], but never bigger than ~560 bp) and can form at multiple sites (max 3) on the same molecule (Figures 3F and 3G). Most structures were resolved when a force >than 20 pN was applied, while others (the one forming a ~350-bp structure) can remain stable until the overstretching transition (up to 60 pN). When the same CENP-B-bound CenX DNA was stretched multiple times, the length of the DNA increased over time, and the discontinuities in the curve disappeared (Figure S3G). This suggested that the secondary DNA structures were resolved during stretching, likely due to the unbinding of CENP-B from the DNA or to the CENP-B/CENP-B dimers going apart.

To further examine the force dependence of CENP-B binding to CenX, we performed force-clamp experiments. In these experiments, the DNA was held at a constant low force (5, 10, or 15 pN) maintained by a feedback loop, while the length of the DNA was measured. When CenX was clamped in the absence of CENP-B, the apparent contour length (calculated with the eWLC model) showed only a single peak at ~1 μm (Figures 3H and S3H; Table S1). Interestingly, contrary to the typical behavior of dsDNA, the exact contour length slightly changed with force from 0.97 μm at 5 pN to 1.02 μm at 15 pN, which presumably is related to the force-dependent dissolution of non-B-form DNA structures such as hairpins. When CenX was incubated with FL-CENP-B just prior to clamping, we frequently observed sudden stepwise elongation of the DNA even at a constant low force, with many peaks observed between 0.75 and 1 μm at 5 pN (Figures 3I, 3J, and S3I). This stepwise elongation presumably originates from opening of several DNA structures formed by CENP-B due to varying distances of the CENP-B-binding sites (Figures 3I and 3J). We never observed structure larger than ~700 bp, as we never detected peaks lower than 0.75 μm, similarly to the steps in the force-distance curves (Figures 3D–3G). The distribution of the apparent contour length changed drastically in relationship to the forces. At 10 pN, most of these peaks disappeared suggesting a lower stability due to the applied force. At this constant force only several peaks around 1 μm remained, with the addition of a unique peak between 0.8 and 0.85 μm that reflects structures of ~350 bp, as observed in the force/distance curves (Figures 3F and 3G). At 15 pN, only the peak around 1 μm remained, indicating that all previous DNA structures had resolved under the impact of the constant applied force. To conclude, force/distance curves on single dsDNA molecules revealed that CENP-B compacts centromeric DNA and promotes formation of secondary structures of a defined size.

CENP-B binding to centromeric DNA forms DNA loops via its dimerization domain

We then used AFM again to investigate which type of secondary structures were present. To this end, we deposited DNA incubated with increasing concentrations of CENP-B (between 4.5 and 18 nM) on mica surfaces with the addition of spermidine to kinetically trap the DNA-protein complexes (Japaridze et al., 2016; Pastré et al., 2006). We then compared the various physical properties of the CenX/CENP-B FL or ΔC complexes with the values obtained for the control samples (Figure 4A). Even at the highest concentration (18 nM), FL-CENP-B did not change the contour length of the DNA compared with DNA alone (Figure 4B; Table S1), indicating that the proteins were neither wrapping nor changing the local twist of the DNA. Importantly, the free radius of gyration of the molecules, a measure describing their average globular size, changed upon binding of FL-CENP-B (150 ± 40 nm for CenX DNA alone compared with 120 ± 35 nm for the highest protein concentration; mean ± SD, Figure 4C; Table S1). The persistence length of CenX also decreased significantly when CENP-B was added (from 52 ± 3 nm in CenX to 35 ± 3 nm in CenX + CENP-B; Table S1). Changes in the free radius of gyration and persistence length were not observed on the pSK control DNA at the same FL-CENP-B protein concentrations and were only moderately detected when the dimerization domain of CENP-B was removed (Figure 4C; Table S1). Altogether, this indicates that CENP-B increases the compaction of centromeric DNA in agreement with the optical tweezers data (Figure 3C).

AFM analysis revealed the presence of DNA loops (Figures 4D, S4A, and S4B). Linear DNA fragments can randomly form DNA loops by self-crossings upon mica surface deposition, especially in the presence of spermidine (Figure S4C). In order to distinguish between such DNA-self crossings and the CENP-B-mediated loops, we measured the frequency of the DNA loops, their size and position relative to the nearest DNA linear end (Figure 4A). The frequency of DNA loops in the CenX DNA increased with increasing concentrations of FL-CENP-B, up to almost double compared with the pSK control (from 36% to 65%, respectively), while this was not observed in the case of the

(F) Bar graphs show the percentage of molecules with >2 DNA loops in the indicated condition. Error bars show the SEM.
(G) Violin plots showing the DNA loop length in the indicated conditions. Horizontal dashed lines marked the expected loop size (highlighted as blue background). Bars represent median, dotted bars quartiles. See Table S1 for details.
(H) Tukey box plot showing the distance from the DNA loop to the nearest linear DNA end. Lines mark the position of the first canonical CENP-B boxes. See Table S1 for details.
For all graphs, the center line marks the median and significance was tested using the Mann-Whitney U test, ns: $0.05 < p$, *: $0.01 < p < 0.05$, **: $0.001 < p < 0.01$, ***: $0.0001 < p < 0.001$, ****: $p < 0.0001$.

CENP-B$^{\Delta C}$ (Figure 4E). At the highest protein concentration, half (50.4%) of all looped DNA structures formed two or three loops (only 17.3% and 22.7% for CenX alone or CenX with CENP-B$^{\Delta C}$, respectively; Figure S4D) with three-looped DNA specifically accumulating (22% for CenX + FL-CENP-B versus 6.9% for CENP-B$^{\Delta C}$) (Figures 4F and S4D).

The DNA loop sizes were also different between the different conditions; loops formed by random DNA deposition were of unspecific sizes in CenX alone (230 ± 120 nm) and similar to the control DNA (Figure 4G; Table S1). In this case, the loop size was unaffected by the addition of the protein to control pSK DNA (215 ± 115 nm and 230 ±115 nm at 0 and 18 nM CENP-B, respectively, mean ± SD) or by adding CENP-B$^{\Delta C}$ to CenX (210 ± 90 nm) (Figure 4G). On the contrary, CENP-B bound to CenX led to smaller loops in a concentration-dependent manner and with a defined size (145 ± 55 nm) (Figure 4G). A loop length that spans between 80 and 160 nm occurs in 60% of the observed molecules at the highest protein concentration (Figure S4E) and matches the distance between two consecutive CENP-B boxes (354–560 bp) (Figures 4G and S2C) and the shortening of the apparent contour length observed by optical tweezers (Figures 3F–3J). Similar changes in the loop size upon CENP-B binding were observed in control DNA that contains five α-sat monomers with two CENP-B boxes positioned 700 bp apart (Figure S4F).

Finally, FL-CENP-B-dependent loops on CenX were mainly positioned between 117 and 185 nm away from the nearest linear DNA end, in agreement with the position of the CENP-B boxes along the DNA molecule (Figure S2C), with loops being preferably formed toward the DNA ends. On the contrary, loops were randomly positioned in CenX without CENP-B (mean of CenX = 207 ± 94 nm versus CenX + CENP-B = 154 ± 48 nm, mean ± SD) (Figures 4H and S4G). In summary, the AFM data indicate that CENP-B compacts centromeric DNA via its dimerization domain by forming ~350–550 bp sized DNA loops, in agreement with optical tweezers experiments.

CENP-B compacts and clusters centromeric chromatin

To assess whether the observed CENP-B-mediated DNA loops reshape centromeric DNA in living cells, we used 3D structured illumination microscopy (3D-SIM) of immunofluorescence (IF)-Fluorescence *in situ* hybridization (FISH) against CENP-C and α-sat. Here, we used pseudodiploid (average of ~44 chromosomes; Dumont et al., 2020) colorectal cancer DLD-1 cells in wild-type and CENP-B-depleted condition (Hoffmann et al., 2020), further rescued with ectopic DOX-inducible FL-CENP-B or CENP-B$^{\Delta C}$ transcribed from an isogenic FRT site. We synchronized and analyzed cells in G1 to minimize variations of centromere DNA due to different phases of the cell cycle. Image analysis on hundreds of centromeres showed that centromeric DNA (visualized with α-sat) is more compacted in wild-type cells compared with CENP-B KO cells where α-sat is more disorganized and dispersed (circularity: 0.91 ± 0.002 in WT versus 0.83 ± 0.003 in KO, mean ± SEM; Figures 5A and 5B; Videos S1 and S2). This decompaction is mirrored as an increase in the total area of centromeric DNA in CENP-B KO cells (0.29 ± 0.007 μm^2 in WT versus 0.42 ± 0.01 μm^2 in KO, mean ± SEM; Figures 5A and 5C; Videos S1 and S2). Importantly, these features are not cell line dependent as they were also observed in RPE-1 cells (a diploid, non-transformed cell line) (Figures 5B, 5C, and S5A) and in mouse fibroblasts depleted for CENP-B (Figures S5B–S5D). Over-expression of FL-CENP-B, but not CENP-B$^{\Delta C}$, further enhanced centromeric DNA compaction (Figures 5A, 5B, and S5E; Videos S3 and S4) reinforcing the role of CENP-B-mediated DNA loops in reshaping centromeric DNA as observed by optical tweezers (Figure 3C) and AFM (Figure 4C). Over-expressed CENP-B did not, however, reduce centromeric area, but rather caused an increase compared with wild-type cells (Figure 5C). This increase derived from clustering of various centromeres in G1-arrested cells as measured by a reduction in the number of distinct visible α-sat foci (Figure 5D). This clustering is dependent on the presence and abundance of CENP-B, but it is only partially dependent on its dimerization domain (Figures 5A and 5D; Videos S1, S2, S3, and S4).

3D-SIM image analysis of IF staining with CENP-C and an anti-centromere antibody (ACA), that recognizes mainly CENP-B, on DLD-1 cells showed that highly expressed FL-CENP-B organized centromeres into hollow, sphere-like structures of ~1–3 μm decorated with several distinct CENP-C foci in a CENP-B concentration-dependent manner (Figures 5E–5H and S5F–S5H; Video S5). Removal of the CENP-B dimerization domain reduced the formation of these structures, their circularity, and, partially, the extent of centromeric clusters (Figures 5E–5H and S5H; Video S6). Changes in centromere organization were also detectable when CENP-B was expressed at its physiological level (Figures S5H–S5K) confirming the IF-FISH data (Figures 5A–5D). Altogether, our results show that CENP-B is capable, via its dimerization domain, to compact and partially cluster centromeric chromatin via the formation of highly ordered centromeric structures.

DNA looping and compaction favor maintenance of centromere position and integrity

We hypothesize that the CENP-B-mediated loops and DNA compaction can provide a favorable environment that contribute to maintain centromere position in interphase and/or centromere integrity during mitosis.

To test the first hypothesis, we took advantage of our recently developed CENP-A/C$^{OFF/ON}$ system that allows us to conditionally abrogate and, subsequently, reactivate the centromere epigenetic marks CENP-A and CENP-C with unprecedented time resolution (Hoffmann et al., 2020). Using this system, we previously demonstrated a key role for CENP-B in preserving centromere position (Hoffmann et al., 2020). Here, we co-depleted and reactivated both CENP-A and CENP-C in cells expressing FL-CENP-B or CENP-B$^{\Delta C}$ (Hoffmann et al., 2020; Figures 6A and S6A). Following depletion (via Auxin, IAA) and reactivation (via IAA removal) of CENP-A/C, cells failed to reload CENP-C in CENP-B downregulated cells (Figures 6B, 6C, S6A, and S6B), as previously shown (Hoffmann et al., 2020). Expression of DOX-inducible FL-CENP-B, but not CENP-B$^{\Delta C}$, restored centromeric CENP-C loading to a similar level as in control condition (no RNAi) (Figures 6B, 6C, S6A, and S6B). These data suggest that the dimerization domain of CENP-B favors centromere maintenance by promoting CENP-C, and subsequently,

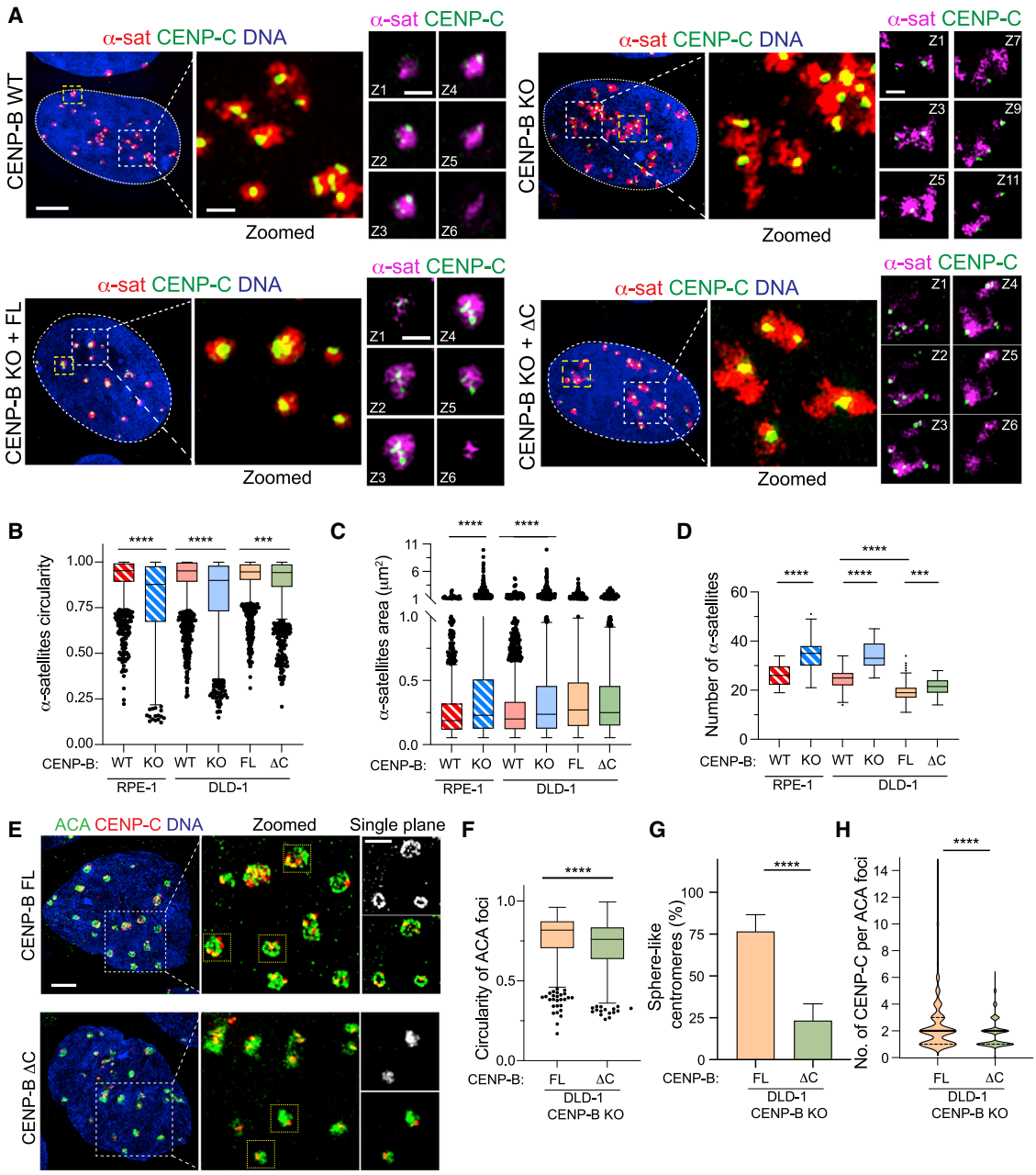


Figure 5. CENP-B compacts centromeric DNA and favors centromere clustering in living cells

(A) Representative 3D-SIM images in the indicated cell lines arrested in G1 with Palbociclib. White square marks a zoomed region, while yellow squares highlight centromeres in single planes (Z = 0.125 μm). To express FL-CENP-B and CENP-B$^{\Delta C}$ Dox was used at 0.1 μg/mL. Scale bars, 5 and 1 μm.

(B–D) Tukey box plots showing circularity (B), area (C), and number (D) of α-sat foci in the indicated cell lines. N ≥ 1,000 centromeres, >40 cells. See Table S1 and Data S1 for details.

(E) Representative 3D-SIM images with the indicated CENP-B constructs. White square marks a zoomed region, while yellow squares highlight centromeres in a single plane. To express FL-CENP-B and CENP-B$^{\Delta C}$ Dox was used at 1 μg/mL. Scale bars, 5 and 2 μm.

(F–H) Tukey box plot, bar graphs and violin plot showing quantification of circularity of ACA foci (F), the type of centromeres (G) and numbers of CENP-C foci per ACA foci (H), respectively, in the indicated cell lines. Error bars in G show the SEM.N ≥ 400 centromeres. See Table S1 for details.

For all graphs the center line marks the median and significance was tested using the Mann-Whitney U test, ns: 0.05 < p, *: 0.01 < p < 0.05, **: 0.001 < p < 0.01, ***: 0.0001 < p < 0.001, ****: p < 0.0001.

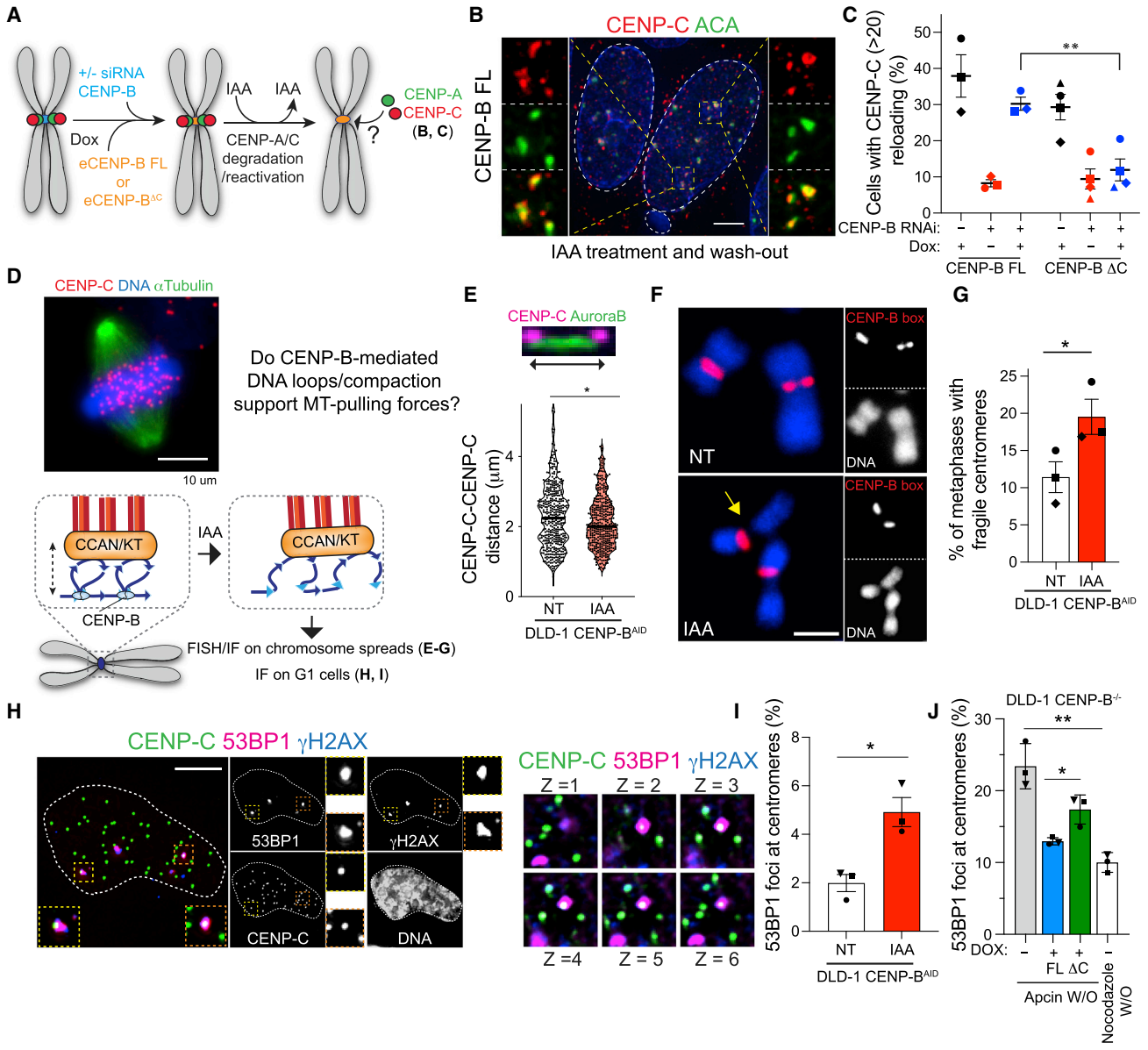


Figure 6. CENP-B-mediated DNA loops and compaction favor centromere position and centromere stability in mitosis

(A) Schematic of the experiment shown in (B). e = ectopic; IAA = Auxin.

(B) Representative immunofluorescence image showing CENP-C and ACA localization after IAA treatment and wash-out. See Figure S6A for details. Yellow squares mark zoomed regions. Scale bars, 5 μm.

(C) Dot blot showing CENP-C reloading in the indicated cell lines and conditions. Each dot represents one experiment with n > 30 cells and each experiment is depicted with a different shape. The bar represents the SEM. Unpaired t test, **p = 0.0052. N = average of 68 cells per experiment. See Table S1 for details.

(D) Schematic representing the aim of the experiment shown in (E–I). See Figure S6D for experimental details.

(E) Violin plot and representative image showing inter-centromere distance between CENP-C dots (magenta) and marked by Aurora B (green). Bars represent the median, unpaired t test, *p = 0.015. N = average of 128 centromeres per experiment. See Table S1 for details.

(F) Representative FISH images of a chromosome with reduced DAPI staining (marked by a yellow arrow) at the centromere, identified with a CENP-B boxfish probe. Scale bars, 3 μm.

(G) Quantification of the percentage of metaphase spreads showing reduced DAPI staining at the centromere. Each dot represents one experiment (depicted with a different shape) with N > 60 cells. The bars represent the SEM. χ^2 test, *p = 0.0343. N = average of 76 cells per experiment. See Table S1 for details.

(H) Representative immunofluorescence images showing CENP-C, γH2AX, and 53BP1. Yellow and orange squares mark zoomed regions. Scale bars, 5 μm.

(I and J) Quantification of the percentage of interphase centromeres showing 53BP1 staining in the indicated cell lines and conditions. Each dot represents one experiment (depicted with a different shape) with an average of >40 cells/experiment. Dox: 1 ng/mL. The bars represent the SEM. Unpaired t test, *p = 0.02, **p = 0.002. N = average of 40 cells per experiment. See Table S1 for details.

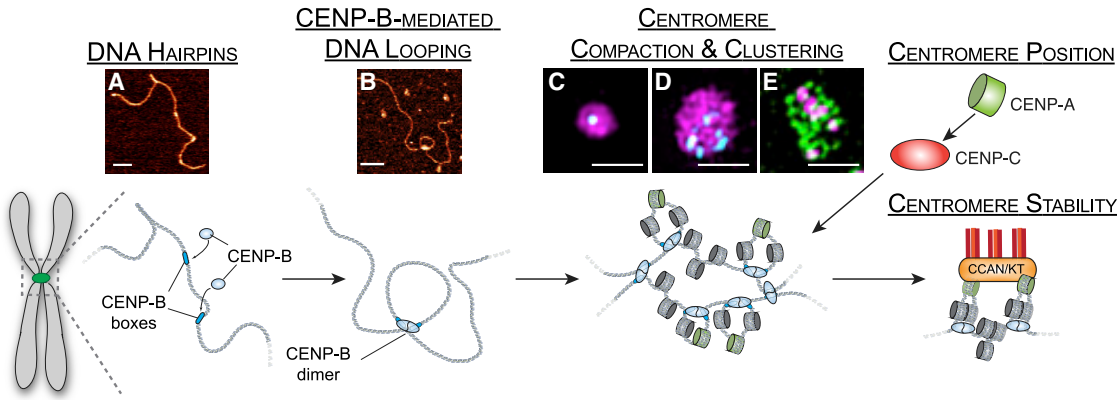


Figure 7. Model of centromeric DNA structure and function via CENP-B-mediated DNA loops

(A–E) Corresponding AFM images show (A) a hairpin of 50 nm (in CenX DNA) and (B) a DNA loop of 165 nm (in CenX + FL-CENP-B) with a scale bars, 100 and 200 nm, respectively. The 3D-SIM images in FL-CENP-B over-expressing cells show (C) compaction of α-sat DNA (magenta) bound by CENP-C (cyan) and two centromeric clusters (D: α-sat DNA in magenta, CENP-C in cyan; E: CENP-B in green, CENP-C in purple). Scale bars, 1 μm.

CENP-A deposition, possibly via the formation of centromeric DNA loops and/or compaction.

We then tested whether these CENP-B-mediated DNA loops could be important during mitosis, as tension develops in this phase in response to external forces. Based on the force dependence of DNA loop formation observed by optical tweezers, we hypothesize that these DNA loops can function as a reservoir of springs to absorb MTs-pulling forces (Lawrimore and Bloom, 2019) and to prevent centromeric breakage (Figure 6D). MT forces have been already shown to promote stretching of centromeres (Cimini et al., 2001; Suzuki et al., 2014; Wan et al., 2009) and, in certain conditions, loss of centromere integrity (Addis Jones et al., 2019; Lera et al., 2019). We set up a protocol to specifically deplete centromeric CENP-B during G2/mitosis (via Auxin-mediated degradation) and to extend MT pulling up to 3 h (via inhibition of the anaphase promoting complex/cyclosome [APC/C]) to exacerbate centromere tension (Figures S6C and S6D). Constant MT pulling on kinetochores led to the recruitment of the mitotic spindle checkpoint protein BUBR1 (Figure S6E), as already observed with the MT-stabilizing drug Taxol (Taylor et al., 2001). Under these conditions and following CENP-B depletion in G2/M, inter-centromere distance was reduced (Figures 6E, S6F, and S6G). This suggests that CENP-B-mediated DNA loops are important for creating a certain distance between sister chromatids, which could serve to absorb MT-pulling forces. Following CENP-B removal, we observed a significant increase in centromere fragility (detected as reduced 4,6-diamidino-2-phenylindole [DAPI] staining) on one or more metaphase chromosomes possibly due to chromosome breakage or decondensation of centromeric regions (Figures 6F and 6G), similarly to what we observed in CENP-A-depleted cells (Giunta et al., 2021). Fragility of centromeres during constant MT pulling and following CENP-B depletion in G2/M was also confirmed by the accumulation at centromeric sites of the DNA-damage markers γH2AX in mitosis and 53BP1 foci in the following G1 (Figures 6H, 6I, and S6H–S6L). Expression of FL-CENP-B significantly rescued accumulation of 53BP1 foci at constitutive CENP-B-depleted centromeres, while expression of CENP-B$^{\Delta C}$ was shown to be less efficient (Figure 6J). Centromeric 53BP1 foci also significantly decreased in CENP-B KO cells released in the absence of prolonged MT pulling (nocodazole release), confirming that prolonged MT-kinetochore tension impacts centromere stability (Figure 6J).

In summary, lack of CENP-B-mediated DNA loops and consequent DNA decompaction and reduced flexibility enhances centromere fragility under MT forces.

DISCUSSION

The results presented in this manuscript change our perspective of centromeres in many ways. We first demonstrate that human centromere DNA folds into non-B-form DNA such as hairpins and more complex folded structures (Figures 1 and 7). Intriguingly, these structures were observed in all the types of centromeric DNA that we analyzed suggesting an intrinsic tendency to form non-B-form DNA. Secondary structures were not only observed in "*in silico-predicted*" non-B-form DNA (as the CENP-B-negative CenY in human and Cen7 in AGM; Kasinathan and Henikoff, 2018) but also found in human centromeres that derived from chromosome X and 3. This finding indicates that secondary structures can form independently of the presence of CENP-B boxes and also contrary to *in silico* predictions, as observed in activated B cells in both minor and major satellites (Kasinathan and Henikoff, 2018; Kouzine et al., 2017). This is likely due to the presence at centromeric DNA of highly repetitive, AT-rich regions that favor the formation of secondary structures such as hairpins. *In vivo*, it is possible that negative supercoiling induced by the RNA polymerase II activity could also facilitate the formation of sequence-specific non-B DNA structures (Talbert and Henikoff, 2020). Centromeric hairpins involve dsDNA interactions, vary in position and length, and can form within the same monomer or two adjacent ones. This suggests that these structures might be highly dynamic and short-lived, a hypothesis also supported by optical tweezers data where we did not observe a stepwise transition between shorter (non-B-form DNA) and longer (B-form DNA) states of the DNA in

both force-distance (Figure 3B) and force-clamp (Figure 3H) experiments.

CENP-B has been predicted to form and/or stabilize small cruciform structures, therefore substituting the spontaneous formation of non-B DNA (Kasinathan and Henikoff, 2018). Here, we have shown that CENP-B is capable of modulating the topology of centromeric DNA by forming structures that are larger than cruciform or hairpin-like structures; single-molecule analysis by AFM and optical tweezers revealed that CENP-B via its DNA-binding and dimerization domains promotes the formation of DNA loops of ~350–560 bp (~115–185 nm) in length between different α-sat monomers (Figures 3, 4, and 7). Interestingly, similarly sized DNA loops were observed during replication of centromeric-DNA-containing BAC in *Xenopus* sp. egg extracts (Aze et al., 2016). Formation of DNA loops via CENP-B's dimerization region consequently promotes compaction of centromeric DNA into sphere-like structures leading to a reduction of total centromeric DNA area in interphase cells (Figures 5 and 7). Similar ring-like organizations were also observed in mitotic kinetochores in *Xenopus* sp. and human cells to promote efficient capture of MTs (Sacristan et al., 2018; Wynne and Funabiki, 2016). Since CENP-B$^{\Delta C}$ still showed a certain degree of chromatin compaction, other domains of CENP-B likely involved in heterochromatin (Otake et al., 2020) and/or transcription (Chen et al., 2021) regulations might also be potentially participating in this process.

Centromere compaction mediated by CENP-B might facilitate bundling of CENP-A nucleosomes that are expected to be spaced between each other. In favor of this hypothesis, high-order centromeric structures of ~100–300 nm containing several CENP-A nucleosomes were already observed (Andronov et al., 2019) where other centromeric proteins such as CENP-C (Melters et al., 2019) and CENP-N (Pentakota et al., 2017; Zhou et al., 2021) could provide additional nucleosome bridging. Here, we show that the CENP-B-mediated compaction of centromeric chromatin in interphase favors CENP-C recruitment and, overall, maintenance of centromere identity by possibly bridging several centromeric nucleosomes creating a special chromatin environment (Figures 6A–6C and 7). An alternative, but not mutually exclusive model, is that secondary structures similar to DNA loops might directly favor CENP-A deposition (Mellone and Fachinetti, 2021), since centromeric DNA was shown to be a preferential site for centromere formation in human artificial chromosomes (HACs) (Ikeno and Hasegawa, 2020), on the Y centromere in human cells (Hoffmann et al., 2020), and in BAC containing centromeric DNA (Aze et al., 2016). In this context, the Holliday junction recognition protein (HJURP) necessary for CENP-A incorporation (Fukagawa and Earnshaw, 2014) may be an important factor due to its ability to recognize cruciform structures on DNA (Kato et al., 2007).

Beside these intrachromosome connections, we report that CENP-B promotes inter-chromosome interactions favoring centromeres clustering with a length of 1–2 μm (Figures 5 and 7). To date, no other factor was clearly implicated in centromere clustering other than condensin II (Hoencamp et al., 2021), and, while centromere clustering is commonly observed in several species, in human, it was mainly reported in non-dividing/quiescent lymphocytes (Muller et al., 2019). The ability of CENP-B to bring together multiple centromeres could have functional relevance to phase-separate certain centromeres to further control DNA metabolism, such as replication or repair of centromeric regions. Further, the possible existence of chromosome-specific patterns within centromeric clusters (Muller et al., 2019) could explain the existence of certain whole-arm chromosomal translocations in pathological conditions such as cancer. To which extent these clusters modulate the different levels of organization of the eukaryotic genome will require further studies. An interesting role could be attributed to DNA methylation (Scelfo and Fachinetti, 2019), since it was proposed to inhibit CENP-B binding (Tanaka et al., 2005b) and CENP-A-chromatin is known to be built on hypomethylated regions (Logsdon et al., 2021; Miga et al., 2020).

Finally, our data have direct implications on the maintenance of genome stability, as centromere breakage and translocation are commonly observed in several types of cancers. Non-B-form DNA could impact replication fork progression leading to double-strand breaks and HORs amplification and evolution (Balzano and Giunta, 2020; Rice, 2019). In this view, centromeres could represent regions with enhanced chromosomal fragility (Barra and Fachinetti, 2018) where CENP-A (Giunta et al., 2021), mismatch repair proteins (Aze et al., 2016), specialized helicases (Li et al., 2018), proteins involved in the homologs recombination pathway (Yilmaz et al., 2021), and other potential factors are necessary to control their stability. While the role of the CENP-B-mediated DNA loops during replication remains to be investigated, we propose that they are important to sustain centromere stability under MT tension. DNA loops could therefore favor the stretching of centromeres/kinetochores (visualized as inter-kinetochore distance) to absorb the tension generated during MT pulling and shortening, which is in the range of several hundreds of pNs (Anjur-Dietrich et al., 2021). This DNA looping mechanism and/or chromatin compaction promoted by CENP-B could facilitate chromosome segregation and maintenance of centromere stiffness, since the absence of such DNA loops causes an increased rigidity and subsequent fragility of centromeric DNA (Figures 6D–6J and 7). A similar spring-like function for centromeres necessary to counteract MT dynamics has been proposed in budding yeast (Lawrimore and Bloom, 2019; Lawrimore et al., 2015). As CENP-B is not ubiquitously present in all vertebrates and CENP-B-deficient centromeres exist even in human (e.g., neocentromeres or the Y chromosome), an important further question is whether there are alternative mechanisms to modulate centromere topology.

In summary, our findings provide a significant step forward in understanding of the basic architecture and organization of human centromeres. While CENP-B was described to play a role in centromere formation (Ohzeki et al., 2020) and maintenance (Hoffmann et al., 2020), heterochromatin establishment (Okada et al., 2007) and kinetochore stability (Dumont et al., 2020; Fachinetti et al., 2015; Hoffmann et al., 2016) via protein-protein interactions, this work provides implications on the importance of DNA topology for centromeric regulation. Our data indicate that organization and regulation of human centromeres resemble those of other repetitive regions of the human genome, such as telomeres and transposable elements, known to form DNA loops via specialized protein complexes and to cluster to adopt a

certain genome architecture (Adam et al., 2019; Cournac et al., 2016).

Limitations of the study

The work is limited to the study of four different types of centromeric DNA molecules (three derived from human and one from *Chlorocebus sabeus*) and three non-centromeric ones with a length of ~2–4 kb, two human cultured cell backgrounds (DLD-1 and RPE-1), and one derived from *Mus musculus* (mouse embryonic fibroblasts). The lack of non-B-form DNA structures outside the analyzed regions or in other centromeric DNA and/or the absence of CENP-B-mediated centromere compaction/clustering in other cell types or in the context of an animal cannot be excluded.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Cell culture
- METHOD DETAILS
  - Gene/protein symbols and ID
  - DNA cloning
  - Cell culture compounds, RNAi and generation of stable cell lines
  - Immunoblotting, immunofluorescence, chromosome spreads and FISH
  - Imaging analysis
  - Protein expression and purification
  - **MBP-CENP-B FL and ΔC purification buffers:**
  - Stoichiometry of the complexes
  - Binding assays
  - Optical tweezers
  - Analysis of optical tweezers data
  - AFM sample preparation
  - AFM imaging
  - Quantification and statistical analysis of AFM images
  - Nanopore experiments
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.molcel.2022.02.032.

ACKNOWLEDGMENTS

The authors would like to thank Erwin J.G. Peterman and Andreas Biebricher (VU University, NL), L. Mirny (MIT, US), and from I. Curie: C. Bartocci, Aurelien Dauphin, R. Gamba, S. Hoffmann, A. Scelfo, S. Herve, and all other Fachinetti team members for helpful suggestions. We would also like to thank A. Straight (Stanford, US) for helpful suggestions, A. Houdusse (I. Curie, FR) for technical suggestions and the usage of biophysics equipment, R. Fujita and H. Kurumizaka (U. Tokyo), C. Francastel (Paris Diderot, FR), and H. Masumoto (Kazusa DNA Research Institut, Japan) for providing reagents, J. Cabanas-Danes, J. Boggers, A. Mukhortava, and all the Lumicks team for helpful suggestions, providing reagents and the hospitality. We also thank the flow cytometry platform, the cell and tissue imaging facility (PICT-IBiSA, member of the French National Research Infrastructure France-BioImaging ANR10-INBS-04), the antibody facility platform, and the sequencing platform at Institut Curie.

D.F. receives salary support from the CNRS. D.F. has received support for this project by Labex « CellnScale», the Institut Curie, the ATIP-Avenir 2015 program, the program «Investissements d'Avenir» launched by the French Government and implemented by ANR with the references ANR-10-LABX-0038 and ANR-10-IDEX-0001-02 PSL , the Emergence grant 2018 from the city of Paris and Fondation ARC pour la Recherche sur Le Cancer PGA1 RF20190208583. C.C. is supported by the Institut Curie EuReCa PhD Programme, Marie Skłodowska-Curie Actions, 847718. C.D. acknowledges supported by the ERC Advanced Grant LoopingDNA (no. 883684) and the Netherlands Organization for Scientific Research (NWO/OCW), as part of the NanoFront and BaSyC programs. A.J. acknowledges support by the Swiss National Science Foundation (grant no. P300P2_177768). G.W. has received funding for this project from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. [883240]) and support by the Netherlands Organisation for Scientific Research (NWO/OCW), as part of the BaSyC program. H.W. would like to thank the Deutsche Forschungsgemeinschaft for financial support (WI 5434/1-1). A.-S.M. is part of the UMR 144-Cell and Tissue Imaging Facility (PICT-IBiSA).

AUTHOR CONTRIBUTIONS

F.C. conceived and performed the *in vitro* protein purification, interaction assay, and the optical tweezers experiment; A.J. conceived, performed, and analyzed the AFM experiments, and A.J. and W.Y. performed the nanopore analysis, both under the supervision of C.D.; H.W. analyzed and performed some of the optical tweezers experiment under the supervision of G.W.; C.K. performed SEC-MALS experiment; S.G. performed EMSA assay and L.V. the MST and protein purification; T.W., M.D., and C.C. performed the experiment in cells; A.-S.M. designed macro for imaging analysis, and L.V. and F.C. analyzed data; D.F. directed the research and conceived the experimental design; D.F., C.D., and G.W. provide financial support. D.F., F.C., A.J., and H.W. made figures and wrote the manuscript. All authors contributed to manuscript editing.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: July 29, 2021
Revised: February 22, 2022
Accepted: February 23, 2022
Published: March 22, 2022

REFERENCES

Adam, N., Degelman, E., Briggs, S., Wazen, R.-M., Colarusso, P., Riabowol, K., and Beattie, T. (2019). Telomere analysis using 3D fluorescence microscopy suggests mammalian telomere clustering in hTERT-immortalized Hs68 fibroblasts. Commun. Biol. *2*, 451.

Addis Jones, O.A., Tiwari, A., Olukoga, T., Herbert, A., and Chan, K.L. (2019). PLK1 facilitates chromosome biorientation by suppressing centromere disintegration driven by BLM-mediated unwinding and spindle pulling. Nat. Commun. *10*, 2861.

Altemose, N., Logsdon, G.A., Bzikadze, A.V., Sidhwani, P., Langley, S.A., Caldas, G.V., Hoyt, S.J., Uralsky, L., Ryabov, F.D., Shew, C.J., et al. (2021). Complete genomic and epigenetic maps of human centromeres. Preprint at bioRxiv. 2021.07.12.452052.

Andronov, L., Ouararhni, K., Stoll, I., Klaholz, B.P., and Hamiche, A. (2019). CENP-A nucleosome clusters form rosette-like structures around HJURP during G1. Nat. Commun. *10*, 4436.

Anjur-Dietrich, M.I., Kelleher, C.P., and Needleman, D.J. (2021). Mechanical mechanisms of chromosome segregation. Cells *10*, 465.

Aze, A., Sannino, V., Soffientini, P., Bachi, A., and Costanzo, V. (2016). Centromeric DNA replication reconstitution reveals DNA loops and ATR checkpoint suppression. Nat. Cell Biol. *18*, 684–691.

Balzano, E., and Giunta, S. (2020). Centromeres under pressure: evolutionary innovation in conflict with conserved function. Genes-Basel *11*, 912.

Barra, V., and Fachinetti, D. (2018). The dark side of centromeres: types, causes and consequences of structural abnormalities implicating centromeric DNA. Nat. Commun. *9*, 4340.

Baumann, C.G., Smith, S.B., Bloomfield, V.A., and Bustamante, C. (1997). Ionic effects on the elasticity of single DNA molecules. Proc Natl Acad Sci U S A *94*, 6185–6190.

Binnig, G., Quate, C.F., and Gerber, C. (1986). Atomic force microscope. Phys. Rev. Lett. *56*, 930–933.

Black, E.M., and Giunta, S. (2018). Repetitive fragile sites: centromere satellite DNA as a source of genome instability in human diseases. Genes-Basel *9*, 615.

Bloom, K., and Costanzo, V. (2017). Centromere structure and function. Prog. Mol. Subcell. Biol. *56*, 515–539.

Bodor, D.L., Rodríguez, M.G., Moreno, N., and Jansen, L.E.T. (2012). Analysis of protein turnover by quantitative SNAP-based pulse-chase imaging. Curr. Protoc. Cell Biol. *55*, 8.8.1–8.8.34.

Casola, C., Hucks, D., and Feschotte, C. (2008). Convergent domestication of pogo-like transposases into centromere-binding proteins in fission yeast and mammals. Mol. Biol. Evol. *25*, 29–41.

Chen, Y., Zhang, Q., Teng, Z., and Liu, H. (2021). Centromeric transcription maintains centromeric cohesion in human cells. J. Cell Biol. *220*, e202008146.

Cimini, D., Howell, B., Maddox, P., Khodjakov, A., Degrassi, F., and Salmon, E.D. (2001). Merotelic kinetochore orientation is a major mechanism of aneuploidy in mitotic mammalian tissue cells. J. Cell Biol. *153*, 517–527.

Cournac, A., Koszul, R., and Mozziconacci, J. (2016). The 3D folding of metazoan genomes correlates with the association of similar repetitive elements. Nucleic Acids Res. *44*, 245–255.

Dumont, M., and Fachinetti, D. (2017). Centromeres and kinetochores, discovering the molecular mechanisms underlying chromosome inheritance. Prog. Mol. Subcell. Biol. *56*, 305–336.

Dumont, M., Gamba, R., Gestraud, P., Klaasen, S., Worrall, J.T., Vries, S.G.D., Boudreau, V., Luypaert, C.S., Maddox, P.S., Lens, S.M., et al. (2020). Human chromosome-specific aneuploidy is influenced by DNA-dependent centromeric features. EMBO J. *39*, 1130–1132.

Earnshaw, W.C., Sullivan, K.F., Machlin, P.S., Cooke, C.A., Kaiser, D.A., Pollard, T.D., Rothfield, N.F., and Cleveland, D.W. (1987). Molecular cloning of cDNA for CENP-B, the major human centromere autoantigen. J. Cell Biol. *104*, 817–829.

Fachinetti, D., Folco, H.D., Nechemia-Arbely, Y., Valente, L.P., Nguyen, K., Wong, A.J., Zhu, Q., Holland, A.J., Desai, A., Jansen, L.E.T., and Cleveland, D.W. (2013). A two-step mechanism for epigenetic specification of centromere identity and function. Nat. Cell Biol. *15*, 1056–1066.

Fachinetti, D., Han, J.S., McMahon, M.A., Ly, P., Abdullah, A., Wong, A.J., and Cleveland, D.W. (2015). DNA sequence-specific binding of CENP-B enhances the fidelity of human centromere function. Dev. Cell *33*, 314–327.

Fukagawa, T., and Earnshaw, W.C. (2014). The centromere: chromatin foundation for the kinetochore machinery. Dev. Cell *30*, 496–508.

Gallego, J., Chou, S.-H., and Reid, B.R. (1997). Centromeric pyrimidine strands fold into an intercalated motif by forming a double hairpin with a Novel T:G:G:T tetrad: solution structure of the d(TCCCGTTTCCA) dimer11Edited. J. Mol. Biol. *273*, 840–856.

Gamba, R., and Fachinetti, D. (2020). From evolution to function: two sides of the same CENP-B coin? Exp. Cell Res. *390*, 111959.

Garavís, M., Méndez-Lago, M., Gabelica, V., Whitehead, S.L., González, C., and Villasante, A. (2015). The structure of an endogenous Drosophila centromere reveals the prevalence of tandemly repeated sequences able to form i-motifs. Sci. Rep. *5*, 13307.

Ghosh, A., and Bansal, M. (2003). A glossary of DNA structures from A to Z. Acta Crystallogr. D Biol. Crystallogr. *59*, 620–626.

Giunta, S., Hervé, S., White, R.R., Wilhelm, T., Dumont, M., Scelfo, A., Gamba, R., Wong, C.K., Rancati, G., Smogorzewska, A., et al. (2021). CENP-A chromatin prevents replication stress at centromeres to avoid structural aneuploidy. Proc. Natl. Acad. Sci. USA *118*, e2015634118.

Grosberg, A.Y., Khokhlov, A.R., and Onuchic, J.N. (1995). Statistical physics of macromolecules. Phys. Today *48*, 92–93.

Gross, P., Laurens, N., Oddershede, L.B., Bockelmann, U., Peterman, E.J.G., and Wuite, G.J.L. (2011). Quantifying how DNA stretches, melts and changes twist under tension. Nat. Phys. *7*, 731–736.

Hasson, D., Panchenko, T., Salimian, K.J., Salman, M.U., Sekulic, N., Alonso, A., Warburton, P.E., and Black, B.E. (2013). The octamer is the major form of CENP-A nucleosomes at human centromeres. Nat. Struct. Mol. Biol. *20*, 687–695.

Heller, I., Hoekstra, T.P., King, G.A., Peterman, E.J.G., and Wuite, G.J.L. (2014). Optical tweezers analysis of DNA–protein complexes. Chem. Rev. *114*, 3087–3119.

Hoekstra, T.P., Depken, M., Lin, S.-N., Cabanas-Danés, J., Gross, P., Dame, R.T., Peterman, E.J.G., and Wuite, G.J.L. (2017). Switching between Exonucleolysis and replication by T7 DNA polymerase ensures high fidelity. Biophys. J. *112*, 575–583.

Hoencamp, C., Dudchenko, O., Elbatsh, A.M.O., Brahmachari, S., Raaijmakers, J.A., van Schaik, T., Sedeño Cacciatore, Á.S., Contessoto, V.G., van Heesbeen, R.G.H.P., van den Broek, B., et al. (2021). 3D genomics across the tree of life reveals condensin II as a determinant of architecture type. Science *372*, 984–989.

Hoffmann, S., Dumont, M., Barra, V., Ly, P., Nechemia-Arbely, Y., McMahon, M.A., Hervé, S., Cleveland, D.W., and Fachinetti, D. (2016). CENP-A is dispensable for mitotic centromere function after initial centromere/kinetochore assembly. Cell Rep. *17*, 2394–2404.

Hoffmann, S., Izquierdo, H.M., Gamba, R., Chardon, F., Dumont, M., Keizer, V., Hervé, S., McNulty, S.M., Sullivan, B.A., Manel, N., and Fachinetti, D. (2020). A genetic memory initiates the epigenetic loop necessary to preserve centromere position. EMBO J. *39*, e105505.

Holland, A.J., Fachinetti, D., Han, J.S., and Cleveland, D.W. (2012). Inducible, reversible system for the rapid and complete degradation of proteins in mammalian cells. Proc Natl Acad Sci U S A *109*. https://doi.org/10.1073/pnas.1216880109.

Hunter, J.D. (2007). Matplotlib: A 2D graphics environment. Comput. Sci. Eng. *9*, 90–95.

Ikeno, M., and Hasegawa, Y. (2020). Applications of bottom-up human artificial chromosomes in cell research and cell engineering. Exp. Cell Res. *390*, 111793.

Japaridze, A., Renevey, S., Sobetzko, P., Stoliar, L., Nasser, W., Dietler, G., and Muskhelishvili, G. (2017). Spatial organization of DNA sequences directs the assembly of bacterial chromatin by a nucleoid-associated protein. J. Biol. Chem. *292*, 7607–7618.

Japaridze, A., Vobornik, D., Lipiec, E., Cerreta, A., Szczerbinski, J., Zenobi, R., and Dietler, G. (2016). Toward an effective control of DNA's submolecular conformation on a surface. Macromolecules *49*, 643–652.

Japaridze, A., Yang, W., Dekker, C., Nasser, W., and Muskhelishvili, G. (2021). DNA sequence-directed cooperation between nucleoid-associated proteins. Iscience *24*, 102408.

Jonstrup, A.T., Thomsen, T., Wang, Y., Knudsen, B.R., Koch, J., and Andersen, A.H. (2008). Hairpin structures formed by alpha satellite DNA of human centromeres are cleaved by human topoisomerase IIα. Nucleic Acids Res. *36*, 6165–6174.

Kasinathan, S., and Henikoff, S. (2018). Non-B-form DNA is enriched at centromeres. Mol. Biol. Evol. *26*, 1301–1314.

Kato, T., Sato, N., Hayama, S., Yamabuki, T., Ito, T., Miyamoto, M., Kondo, S., Nakamura, Y., and Daigo, Y. (2007). Activation of Holliday junction recognizing protein involved in the chromosomal stability and immortality of cancer cells. Cancer Res. *67*, 8544–8553.

Kouzine, F., Wojtowicz, D., Baranello, L., Yamane, A., Nelson, S., Resch, W., Kieffer-Kwon, K.-R., Benham, C.J., Casellas, R., Przytycka, T.M., et al. (2017). Permanganate/S1 nuclease footprinting reveals non-B DNA structures with regulatory potential across a mammalian genome. Cell Syst. *4*, 344–356.e7.

Kumar Sharma, R.K., Agrawal, I., Dai, L., Doyle, P.S., and Garaj, S. (2019). Complex DNA knots detected with a nanopore sensor. Nat. Commun. *10*, 4473.

de Lange, T. (1992). Human telomeres are attached to the nuclear matrix. EMBO J. *11*, 717–724.

Lawrimore, J., and Bloom, K. (2019). The regulation of chromosome segregation via centromere loops. Crit. Rev. Biochem. Mol. Biol. *54*, 352–370.

Lawrimore, J., Vasquez, P.A., Falvo, M.R., Taylor, R.M., Vicci, L., Yeh, E., Forest, M.G., and Bloom, K. (2015). DNA loops generate intracentromere tension in mitosis. J. Cell Biol. *210*, 553–564.

Lera, R.F., Norman, R.X., Dumont, M., Dennee, A., Koob, J.M., Fachinetti, D., and Burkard, M.E. (2019). Plk1 protects kinetochore–centromere architecture against microtubule pulling forces. EMBO Rep. *20*, 805–816.

Li, H., Xiao, J., Li, J., Lu, L., Feng, S., and Dröge, P. (2009). Human genomic Z-DNA segments probed by the Zα domain of ADAR1. Nucleic Acids Res. *37*, 2737–2746.

Li, Z., Liu, B., Jin, W., Wu, X., Zhou, M., Liu, V.W., Goel, A., Shen, Z., Zheng, L., and Shen, B. (2018). hDNA2 nuclease/helicase promotes centromeric DNA replication and genome stability. EMBO J. *37*, e96729.

Logsdon, G.A., Vollger, M.R., Hsieh, P., Mao, Y., Liskovykh, M.A., Koren, S., Nurk, S., Mercuri, L., Dishuck, P.C., Rhie, A., et al. (2021). The structure, function and evolution of a complete human chromosome 8. Nature *593*, 101–107.

Lowary, P.T., and Widom, J. (1998). New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning11Edited. J. Mol. Biol. *276*, 19–42.

Masumoto, H., Masukata, H., Muro, Y., Nozaki, N., and Okazaki, T. (1989). A human centromere antigen (CENP-B) interacts with a short specific sequence in alphoid DNA, a human centromeric satellite. J. Cell Biol. *109*, 1963–1973.

Mellone, B.G., and Fachinetti, D. (2021). Diverse mechanisms of centromere specification. Curr. Biol. *31*, R1491–R1504.

Melters, D.P., Pitman, M., Rakshit, T., Dimitriadis, E.K., Bui, M., Papoian, G.A., and Dalal, Y. (2019). Intrinsic elasticity of nucleosomes is encoded by histone variants and calibrated by their binding partners. Proc. Natl. Acad. Sci. USA *116*, 24066–24074.

Miga, K.H., Koren, S., Rhie, A., Vollger, M.R., Gershman, A., Bzikadze, A., Brooks, S., Howe, E., Porubsky, D., Logsdon, G.A., et al. (2020). Telomere-to-telomere assembly of a complete human X chromosome. Nature *585*, 79–84.

Mikhaylov, A., Sekatskii, S.K., and Dietler, G. (2013). DNA trace: a comprehensive software for polymer image processing. J. Adv. Microsc. Res. *8*, 241–245.

Mirkin, S.M. (2007). Expandable DNA repeats and human disease. Nature *447*, 932–940.

Muller, H., Gil, J., and Drinnenberg, I.A. (2019). The impact of centromeres on spatial genome architecture. Trends Genet. *35*, 565–578.

Muro, Y., Masumoto, H., Yoda, K., Nozaki, N., Ohashi, M., and Okazaki, T. (1992). Centromere protein B assembles human centromeric alpha-satellite DNA at the 17-bp sequence, CENP-B box. J. Cell Biol. *116*, 585–596.

Nečas, D., and Klapetek, P. (2012). Gwyddion: an open-source software for SPM data analysis. Open Phys. *10*, 181–188.

Nechemia-Arbely, Y., Fachinetti, D., Miga, K.H., Sekulic, N., Soni, G.V., Kim, D.H., Wong, A.K., Lee, A.Y., Nguyen, K., Dekker, C., et al. (2017). Human centromeric CENP-A chromatin is a homotypic, octameric nucleosome at all cell cycle points. J. Cell Biol. *216*, 607–621.

Nonin-Lecomte, S., and Leroy, J.L. (2001). Structure of a C-rich strand fragment of the human centromeric satellite III: a pH-dependent intercalation topology 1. J. Mol. Biol. *309*, 491–506.

Odijk, T. (1995). Stiff chains and filaments under tension. Macromolecules *28*, 7016–7018.

Ohzeki, J.I., Otake, K., and Masumoto, H. (2020). Human artificial chromosome: chromatin assembly mechanisms and CENP-B. Exp. Cell Res. *389*, 111900.

Okada, T., Ohzeki, J., Nakano, M., Yoda, K., Brinkley, W.R., Larionov, V., and Masumoto, H. (2007). CENP-B controls centromere formation depending on the chromatin context. Cell *131*, 1287–1300.

Otake, K., Ohzeki, J.I., Shono, N., Kugou, K., Okazaki, K., Nagase, T., Yamakawa, H., Kouprina, N., Larionov, V., Kimura, H., et al. (2020). CENP-B creates alternative epigenetic chromatin states permissive for CENP-A or heterochromatin assembly. J. Cell Sci. *133*, jcs243303.

Pastré, D., Hamon, L., Landousy, F., Sorel, I., David, M.O., Zozime, A., Le Cam, E.L., and Piétrement, O. (2006). Anionic polyelectrolyte adsorption on mica mediated by multivalent cations: a solution to DNA imaging by atomic force microscopy under high ionic strengths. Langmuir *22*, 6651–6660.

Pentakota, S., Zhou, K., Smith, C., Maffini, S., Petrovic, A., Morgan, G.P., Weir, J.R., Vetter, I.R., Musacchio, A., and Luger, K. (2017). Decoding the centromeric nucleosome through CENP-N. Elife *6*, e33442.

Plesa, C., and Dekker, C. (2015). Data analysis methods for solid-state nanopores. Nanotechnology *26*, 084003.

Plesa, C., Verschueren, D., Pud, S., Torre, J. van der, Ruitenberg, J.W., Witteveen, M.J., Jonsson, M.P., Grosberg, A.Y., Rabin, Y., and Dekker, C. (2016). Direct observation of DNA knots using a solid-state nanopore. Nat. Nanotechnol. *11*, 1093–1097.

Rice, W.R. (2019). A game of thrones at human centromeres II. A new molecular/evolutionary model. Preprint at bioRxiv. https://doi.org/10.1101/731471.

Rubinstein, M., and Colby, R.H. (2003). Polymer Physics (Orford University Press).

Sacristan, C., Ahmad, M.U.D., Keller, J., Fermie, J., Groenewold, V., Tromer, E., Fish, A., Melero, R., Carazo, J.M., Klumperman, J., et al. (2018). Dynamic kinetochore size regulation promotes microtubule capture and chromosome biorientation in mitosis. Nat. Cell Biol. *20*, 800–810.

Scelfo, A., and Fachinetti, D. (2019). Keeping the centromere under control: a promising role for DNA methylation. Cells *8*, 912.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682.

Skene, P.J., and Henikoff, S. (2015). A simple method for generating high-resolution maps of genome-wide protein binding. Elife *4*, e09225.

Smith, S.B., Cui, Y., and Bustamante, C. (1996). Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules. Science *271*, 795–799.

Storm, A.J., Chen, J.H., Zandbergen, H.W., and Dekker, C. (2005). Translocation of double-strand DNA through a silicon oxide nanopore. Phys. Rev. E Stat. Nonlin. Soft Matter Phys. *71*, 051903.

Suzuki, A., Badger, B.L., Wan, X., DeLuca, J.G., and Salmon, E.D. (2014). The architecture of CCAN proteins creates a structural integrity to resist spindle forces and achieve proper intrakinetochore stretch. Dev. Cell *30*, 717–730.

Talbert, P.B., and Henikoff, S. (2020). What makes a centromere? Exp. Cell Res. *389*, 111895.

Tanaka, Y., Kurumizaka, H., and Yokoyama, S. (2005b). CpG methylation of the CENP-B box reduces human CENP-B binding. FEBS J. *272*, 282–289.

Tanaka, Y., Nureki, O., Kurumizaka, H., Fukai, S., Kawaguchi, S., Ikuta, M., Iwahara, J., Okazaki, T., and Yokoyama, S. (2001). Crystal structure of the CENP-B protein-DNA complex: the DNA-binding domains of CENP-B induce kinks in the CENP-B box DNA. EMBO J. *20*, 6612–6618.

Tanaka, Y., Tachiwana, H., Yoda, K., Masumoto, H., Okazaki, T., Kurumizaka, H., and Yokoyama, S. (2005a). Human centromere protein B induces translational positioning of nucleosomes on alpha-satellite sequences. J. Biol. Chem. *280*, 41609–41618.

Tawaramoto, M.S., Park, S.-Y., Tanaka, Y., Nureki, O., Kurumizaka, H., and Yokoyama, S. (2003). Crystal structure of the human centromere protein B (CENP-B) dimerization domain at 1.65-Å resolution. J. Biol. Chem. *278*, 51454–51461.

Taylor, S.S., Hussein, D., Wang, Y., Elderkin, S., and Morrow, C.J. (2001). Kinetochore localisation and phosphorylation of the mitotic checkpoint components Bub1 and BubR1 are differentially regulated by spindle events in human cells. J. Cell Sci. *114*, 4385–4395.

Wan, X., O'Quinn, R.P., Pierce, H.L., Joglekar, A.P., Gall, W.E., DeLuca, J.G., Carroll, C.W., Liu, S.T., Yen, T.J., McEwen, B.F., et al. (2009). Protein architecture of the human kinetochore microtubule attachment site. Cell *137*, 672–684.

Witz, G., Rechendorff, K., Adamcik, J., and Dietler, G. (2008). Conformation of circular DNA in two dimensions. Phys. Rev. Lett. *101*, 148103.

Wu, J.C., and Manuelidis, L. (1980). Sequence definition and organization of a human repeated DNA. J. Mol. Biol. *142*, 363–386.

Wynne, D.J., and Funabiki, H. (2016). Heterogeneous architecture of vertebrate kinetochores revealed by three-dimensional superresolution fluorescence microscopy. Mol. Biol. Cell *27*, 3395–3404.

Yilmaz, D., Furst, A., Meaburn, K., Lezaja, A., Wen, Y., Altmeyer, M., Reina-San-Martin, B., and Soutoglou, E. (2021). Activation of homologous recombination in G1 preserves centromeric integrity. Nature *600*, 748–753.

Yoda, K., Ando, S., Okuda, A., Kikuchi, A., and Okazaki, T. (1998). In vitro assembly of the CENP-B/α-satellite DNA/core histone complex: CENP-B causes nucleosome positioning. Genes Cells *3*, 533–548.

Yoda, K., Kitagawa, K., Masumoto, H., Muro, Y., and Okazaki, T. (1992). A human centromere protein, CENP-B, has a DNA binding domain containing four potential alpha helices at the NH2 terminus, which is separable from dimerizing activity. J. Cell Biol. *119*, 1413–1427.

Zhou, K., Gebala, M., Woods, D., Sundararajan, K., Edwards, G., Krzizike, D., Wereszczynski, J., Straight, A.F., and Luger, K. (2021). CENP-N promotes the compaction of centromeric chromatin. Preprint at bioRxiv. https://doi.org/10.1101/2021.06.14.448351.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal anti-CENP-B	Abcam	Cat# ab25734; RRID:AB_726801
Guinea pig polyclonal anti-CENP-C	Clinisiences	Cat# PD030; AB_10693556
Human polyclonal anti-centromere antisera (ACA)	Antibodies Incorporated	Cat# 15-235-0001; RRID: AB_2797146
Rabbit polyclonal anti-CENP-B	Santa Cruz Biotechnology	Cat# sc-22788, RRID: AB_2078775
Mouse monoclonal anti-αTubulin (clone DM1A)	Sigma-Aldrich	Cat# T9026; RRID: AB_477593
Rabbit polyclonal anti-AuroraB	Abcam	Cat# ab2254; AB_302923
Mouse monoclonal anti-Vinculin (clone hVIN-1)	Sigma-Aldrich	Cat# V9264; RRID: AB_10603627
Rabbit polyclonal 53BP1	Novus-biologicals	Cat# NB100-304; AB_10003037
Mouse γH2AX	Millipore/Upstate	Cat# JBW301; AB_568825
Donkey polyclonal anti-mouse - Cy3	Jackson ImmunoResearch Laboratories	Cat# 715-165-151; RRID: AB_2315777
Donkey polyclonal anti-rabbit - Cy5	Jackson ImmunoResearch Laboratories	Cat# 711-175-152; RRID:AB_2340607
Donkey polyclonal anti-mouse - Alexa Fluor® 488	Jackson ImmunoResearch Laboratories	Cat# 715-545-150; RRID: AB_2340846
Donkey polyclonal anti-rabbit - Cy3	Jackson ImmunoResearch Laboratories	Cat# 711-165-152; RRID: AB_2307443
Donkey polyclonal anti-guinea pig - Alexa Fluor® 647	Jackson ImmunoResearch Laboratories	Cat# 706-605-148; RRID: AB_2340476
Sheep anti-mouse - HRP	GE Healthcare	Cat# NA931l; RRID: AB_772212
Donkey anti-rabbit – HRP	GE Healthcare	Cat# NA934; RRID: AB_772206
Bacterial strains		
NEB® 5-alpha Competent E. coli	New England Biolabs	Cat# C2987H
BL21-AI™ One Shot™ Chemically Competent E. coli	Invitrogen	Cat# C607003
Sf9 cells	Gibco	Cat# 11496015
High Five™ Cells	Gibco	Cat# B85502
MAX Efficiency™ DH10Bac™ Competent Cells	Thermo Fischer Scientific	Cat# 10361012
Chemicals, peptides, and recombinant proteins		
Indole-3-acetic acid (IAA) sodium salt	Sigma-Aldrich	Cat# I5148; CAS: 6505-45-9
Polybrene®	Santa Cruz Biotechnology	Cat# sc-134220; CAS 28728-55-4
VectaShield	Vector Laboratories	Cat# H-1000; RRID: AB_2336789
Invitrogen™ PureLink™ HiPure Plasmid Filter Midiprep Kit	Invitrogen	Cat# 210014
QIAGEN Plasmid Plus Kit	QIAGEN	Cat# 12943
NucleoSpin Gel and PCR Clean-up, Mini kit	Macherey-Nagel	Cat# 740609.50
Q5® High-Fidelity DNA Polymerase	New England Biolabs	Cat# M0493L
dNTP set	Invitrogen	Cat# 10297117
Sf-900™ II SFM medium	Gibco	Cat# 10902088
Insect-XPRESS™ Protein-free Insect Cell Medium	Lonza	Cat# BELN12-730Q
Any kD™ Mini-PROTEAN® TGX™ Precast Protein Gels	BioRAD	Cat# 4569036
InstantBlue® Coomassie Protein Stain	Euromedex	Cat# EX-ISB1L
Agarose	Life Technologies	Cat# 16500500
GelGreen® Nucleic Acid Gel Stain	Biotium	Cat# 41005
Tris Acetate EDTA (TAE) buffer	EUROMEDEX	Cat# EU0201-A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Tris Borate EDTA (TBE) buffer	EUROMEDEX	Cat# ET020-A
Sodium Dodecyl Sulfate (SDS)	EUROMEDEX	Cat# EU0660
Glycine	Thermo Fischer Scientific	Cat# 15665690
2xYT medium	Thermo Fischer Scientific	Cat# 11738892
Kanamycin	EUROMEDEX	Cat# UK0015-B
Ampicillin	EUROMEDEX	Cat# EU0400-A
Isopropyl-β-D-1-thiogalactopyranoside (IPTG)	EUROMEDEX	Cat# EU0008-A
L-(+)-Arabinose	Sigma-Aldrich	Cat# A3256-100G
4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)	EUROMEDEX	Cat# 10-110-A
NaCl	Thermo Fischer Scientific	Cat# 16623410
Dithiothreitol (DTT)	EUROMEDEX	Cat# EU0006-B
Complete Protease Inhibitor Cocktail tab	Sigma-Aldrich	Cat# 1187358001
Phenylmethanesulfonyl fluoride (PMSF)	Sigma-Aldrich	Cat# P7626
D-(+)-Maltose	Sigma-Aldrich	Cat# M9171
Glycerol	Thermo Fischer Scientific	Cat# 15665690
ß-Mercaptoethanol	Life Technologies	Cat# 31350010
Imidazole	Sigma-Aldrich	Cat# 56750
KCl	Thermo Fischer Scientific	Cat# 15691160
MgCl2	Sigma-Aldrich	Cat# 13152
NiSO4	Sigma-Aldrich	Cat# 227676
NaN3	Sigma-Aldrich	Cat# 71289
Tween20	Sigma-Aldrich	Cat# P9416
SYBR Safe DNA Gel Stain	Thermo Fischer Scientific	Cat# S33102
1 kb DNA Ladder	New England Biolabs (NEB)	Cat# N3232S
L-Glutamine	Thermo Fischer Scientific	Cat# 25030149
MBP-CENP-B FL	This study	N/A
MBP-CENP-B ΔC	This study	N/A
CENP-B ΔC-6xHis-EGFP	This study	N/A
Chromatography columns 5ml MBPTrap™ HP	GE healthcare	Cat# 28-9187-80
Chromatography columns HiLoad® 16/600 Superdex® 200 pg	GE healthcare	Cat# 28-9893-35
Chromatography columns 5ml HisTrap™ Fast Flow	GE healthcare	Cat# 17-5255-01
Chromatography columns 1ml HiTrap® Capto™ Q ImpRes	GE healthcare	Cat# 17-5470-51
Chromatography columns Superdex® 200 10/300 GL	GE healthcare	Cat# 17-5175-01
Deposited data		
Mendeley dataset	https://doi.org/10.17632/cs9v2zvktv.1	
Experimental models: Cell lines		
Human: hTERT RPE-1 $^{EYFP-AID/EYFP-AID}$CENP-A +OsTIR19Myc-Puro	(Hoffmann et al., 2016)	N/A
Human: hTERT RPE-1 $^{EYFP-AID/EYFP-AID}$CENP-A CENP-B KO +OsTIR19Myc-Puro	(Fachinetti et al., 2015)	N/A
Human: DLD-1 Flp-In T-Rex +OsTIR19Myc-Puro	(Holland et al., 2012)	N/A
Human: DLD-1 CENP-B$^{AID-EYFP}$ +OsTIR19Myc-Puro	(Hoffmann et al., 2016)	N/A
Human: DLD-1 CENP-B KO Flp-In T-Rex +OsTIR19Myc-Puro	(Hoffmann et al., 2020)	N/A
Human: DLD-1 CENP-B KO + FRT::TO-CENP-B(fl)-mCherry +OsTIR19Myc-Puro	(Hoffmann et al., 2020)	N/A
Human: DLD-1 CENP-B KO + FRT::TO-CENP-B(ΔC) +OsTIR19Myc-Puro	This paper	N/A
Human: hTERT DLD-1 $^{EYFP-AID/-}$CENP-A $^{mcherry-AID/AID}$CENP-C +OsTIR19Myc-Puro	(Hoffmann et al., 2020)	N/A

(*Continued on next page*)

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
Mouse: Embryonic Fibroblast	(Okada et al., 2007)	N/A
Mouse: Embryonic Fibroblast CENP-B KO	(Okada et al., 2007)	N/A
Recombinant DNA		
CenX	This paper	N/A
CenY	This paper	N/A
Cen3	This paper	N/A
AGM	This paper	N/A
pSty11	(De Lange, 1992)	Addgene # 12401
601	Gift from H. Kurumizaka	N/A
pSK	pBluescript	N/A
pKYB1	(Hoekstra et al., 2017)	N/A
Software and algorithms		
Fiji	(Schindelin et al., 2012)	https://imagej.net/Fiji
CRaQ v1.12	(Bodor et al., 2012)	http://facilities.igc.gulbenkian.pt/microscopy/macros/CRaQ_v1.12.ijm
Centromere intensity	(Giunta et al., 2021)	Available upon request
Centromere circularity, area and cluster	This study	Included as supplemental information
GraphPad Prism	www.graphpad.com	RRID:SCR_002798
Python	https://www.python.org	3.7.9
Scipy	https://scipy.org	1.6.0
Pylake	https://lumicks-pylake.readthedocs.io/en/stable/	0.8.1
Imaris	https://imaris.oxinst.com/microscopy/software	N/A

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Daniele Fachinetti (daniele.fachinetti@curie.fr).

Materials availability

Plasmids and cell lines generated in this study will be provided upon request to the lead contact.

Data and code availability

- Data reported in this paper will be shared by the lead contact upon request.
- All original code is available in this paper's supplemental information.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.
- Raw data of main images were deposited on Mendeley at https://doi.org/10.17632/cs9v2zvktv.1

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell line name	Type	Species	Original Source
hTERT RPE-1	Retinal pigment epithelial	Human	ATTC
DLD-1	Colon Cancer	Human	ATTC
MEFs	Embryonic Fibroblast	Mouse	Okada et al., 2007

Cell culture

Flp-In TREx DLD-1 cells and Mouse Embryonic Fibroblasts (MEFs a kind gift from Hiroshi Masumoto; Okada et al., 2007) were cultivated at 37°C in Dulbecco's modified essential medium (DMEM) medium containing 10% tetracycline free Fetal Bovine Serum (FBS,

Pan Biotech) at 5% CO2 atmosphere. hTERT RPE-1 cells were maintained in DMEM:F12 (GIBCO) medium containing 10% heat inactivated FBS (biosera), 0.123% sodium bicarbonate and 2 mM L-glutamine.

METHOD DETAILS

Gene/protein symbols and ID

Input	Symbol	Approved name	HGNC ID	Location
CENP-B	CENPB	centromere protein B	HGNC:1852	20p13
CENP-A	CENPA	centromere protein A	HGNC:1851	2p23.3
CENP-C	CENPC	centromere protein C	HGNC:1854	4q13.2
BUBR1	BUB1B	BUB1 mitotic checkpoint serine/threonine kinase B	HGNC:1149	15q15.1
γH2AX	H2AX	H2A.X variant histone	HGNC:4739	11q23.3
53BP1	TP53BP1	tumor protein p53 binding protein 1	HGNC:11999	15q15.3
CENPN	CENPN	centromere protein N	HGNC:30873	16q23.2
HJURP	HJURP	Holliday junction recognition protein	HGNC:25444	2q37.1

DNA cloning

The open reading frame of the *Escherichia coli* maltose/maltodextrin-binding periplasmic protein (MBP) was inserted in fusion at the N-terminal of the open reading frame of *CENPB* FL or ΔC ([1-538]) in pET30 vector (from Fachinetti et al., 2015). The open reading frame of *CENPB* [1-538] was inserted in fusion at the N-terminal of His-*EGFP* in pFastBac plasmid to express CENP-B$^{\Delta C}$.

The 57 bp DNA was obtained by annealing of two synthetic single strand DNA synthetized by Eurogentec and IDT. One of the two strands is labelled with an Alexa Fluor 546.

3 Kb DNA derived from PCR amplification of centromeres of human chromosome X (from hg38: 60,070,156-60,073,157), human chromosome Y (from hg38: 10,329,372-10,332,268), human chromosome 3 (from hg38: 92545021-92548081) and chromosome 7 of African Green Monkeys (from *Chlorocebus sabeus* 1.1: 16899047-16901965) were synthetized from GENEWIZ and cloned into plasmid pUC57. pSty11 containing 800 bp of Telomeric DNA (De Lange, 1992) was a gift from Titia de Lange (Addgene plasmid # 12401; http://n2t.net/addgene:12401). DNA containing 12 tandem repeats of 208 bp Widom 601 sequence (Lowary and Widom, 1998) was a kind gift from Kurumizaka Laboratory. CenX-300, 500 or 700 containing 3, 4 or 5 of α-sat monomers derived from CenX were purchased by IDT and cloned into pBluescript pSK plasmid. pKYB1 DNA was synthesized as described earlier (Hoekstra et al., 2017).

Cell culture compounds, RNAi and generation of stable cell lines

IAA (I5148; Sigma) dissolved in ddH2O was used at 500 μM, doxycycline (from 1ng/ml to 1 μg/ml, as indicated), Palbociclib (1 μM), 9 μM RO-3306, 12 μM proTAME (R&D Systems), 25 μM Apcin (Sigma), Nocodazole (200 ng/ml). IAA was washed-out three times using culture medium.

Lipofectamine RNAiMax (Invitrogen) was used to introduce siRNAs as described previously (Fachinetti et al., 2013). siRNAs SMARTpool against Luciferase and CENP-B were purchased from Dharmacon.

The FRT/Flp-in system was used to generate stable cell lines as described previously (Fachinetti et al., 2013). Briefly, we integrated cDNA of FL-*CENPB*$^{EYFP\text{-}AID}$, FL-*CENPB*^{mcherry} and *CENP-B*$^{\Delta C}$ into a pcDNA5/FRT plasmid harboring a promoter and start-codon lacking hygromycin resistance cassette and co-transfected this plasmid with a pOG44 plasmid (Flp-recombinase) in a 9:1 ratio into Flp-In TREx DLD-1 WT or CENP-B KO cells using FugeneHD (Promega). Correct integration of the insert at the isogenic FRT site gives rise to hygromycin resistance. After selection (0.3 mg/ml hygromycin, Invitrogen) single clones were isolated and tested for successful integration by immunofluorescence microscopy.

Immunoblotting, immunofluorescence, chromosome spreads and FISH

For immunoblotting, cell pellets were suspended in protein sample buffer and samples were separated by SDS-PAGE, transferred onto nitrocellulose membranes (Bio-Rad), and revealed with the following antibodies: CENP-B (1:1,000) and Vinculin (1:2,000). For immunofluorescence, cells were fixed in 4% formaldehyde 0.1% Triton X-100 at room temperature for 10 min or in methanol at -20°C for 15 min and subsequently in a blocking buffer (2.5% FBS (v/v), 0.2 M Glycine, 0,1% triton X-100 (v/v) in PBS) for 30 min at room temperature. Primary antibodies were incubated in blocking buffer for 1 h at room temperature. The following antibodies were used: CENP-C (1:1,000), CENP-B (1:1,000), ACA (1:500), 53BP1 (1:500), αTubulin (1:2000) and γH2AX (1:1000). All secondary antibodies used for immunofluorescence microscopy were purchased from Jackson Immuno Research and used at 1:500. After immunostaining cells were DAPI counterstained and mounted using anti-fading reagent (Life technologies).

Fluorescence in situ hybridization (FISH) was performed as follows. Cells were trapped in mitosis with a mix of 12 μM proTAME and 25 μM Apcin for 3h. Afterwards cells were harvested and swollen in hypotonic solution (75 mM KCl) before fixation with acetic acid/ethanol mix (1:3) and spread onto slides. Coverslips were rinsed with 80 % ethanol, air-dried and probes against CENP-B Box Cy3 probe (PNA, Bio) (1:300 of 50 μM stock) or α-sat strep Alexa 594-biotin (1:100 of 100 μM stock) (Giunta et al., 2021) were applied. Probes and samples were denaturated by heating (65 °C, 2 min) followed by a 2h incubation at 37 °C. Then slides were washed with 0.4X SSC at 72 °C for 2 min, followed by a 30 sec wash in 2X SCC, 0.05 % Tween 20 at RT. Slides were rinsed with PBS before mounting in ProLong Gold antifade reagent with 4,6-diamidino-2-phenylindole (DAPI, P36935, Invitrogen) and imaged with an epi-fluorescence microscope (Leica DM6000) (100 x objective). FISH on MEFs was performed as described above on interphase cells fixed in Carnoy's solution (methanol/acetic acid 3:1), using a biotinylated probe targeting minor satellite sequence (1:100 from a 100 uM stock) and a denaturation step at 72°C for 2 min. Before mounting, slides were incubated with 1:1000 Streptavidin-Alexa Fluor™ 647 conjugate (stock at 1mg/mL, Invitrogen™), washed three times with PBS 1X and mounted.

For IF-FISH, following IF protocol, the cells were post-fixed in 2% formaldehyde for 10 mins prior FISH protocol. For denaturation, coverslips were incubated at 75°C for 2 mins and hybridization was done overnight at 37°C.

Images of DAPI counterstained and immuno-/FISH-stained cells were collected using a Deltavision Core system (Applied Precision) with 100X Olympus UPlanSApo 100 oil-immersion objective (NA 1.4), 250W Xenon light source equipped with a Photometrics CoolSNAP_HQ2 Camera. ~ 3 μm Z-stacks were acquired (Z step size: 0.2 μm).

Structured illumination microscopy was performed using an OMX 3D-SIM system (DeltaVision) with 100X Olympus UPlanSApo 100 oil-immersion objective (NA 1.4). Signal from all channels were realigned using fluorescent beads. Z-stacks were acquired (Z step size: 0.125 μm).

Imaging analysis

An ImageJ macro (Data S1) was written to compute the different parameters of the α-satellites. First, all alpha-satellites were detected in their respective channel. We distinguished the structures from the background using an automatic threshold (Otsu method), after a pre-processing made of a background subtraction and a median filtering. Alpha-satellite area and circularity [circularity = 4pi(area/perimeter^2) a value close to 1 will define a perfect circle, a value close to 0 will describe a more elongated structure] were obtained by the Analyse Particles function. In a second step, the nuclei were detected in the respective channel, using an automatic threshold (Huang method, after a pre-processing made of background subtraction/ Gaussian blur/unsharp mask). The macro offers the possibility to manually adjust each nucleus contour in order to provide a more accurate result. Finally, the macro computed the number of alpha-satellites in each nucleus as well as the percentage of area covered by those structures within each nucleus (that is the sum of the areas of all alpha-satellites within the nucleus divided by the nucleus area).

Protein expression and purification

FL MBP-CENP-B and MBP-CENP-B$^{\Delta C}$ were expressed in BL21-AI™ (Invitrogen life technologies, Carlsbad, CA, USA). Bacteria were grown at 37°C in 2xYT medium supplemented with Kanamycin (50 μg/ml) until OD600nm between 0.6 and 0.8. Recombinant protein expression was induced by addition of 0.5 mM Isopropyl β- D -1-thiogalactopyranoside (IPTG) and 0.2% L-(+)-Arabinose. Cell cultures were incubated overnight at 20°C and then harvested by centrifugation 4500 g for 30 min. CENP-B ΔC-6xHis-EGFP was expressed in Hi-Five insect cells using the Bac-to-Bac expression system (Invitrogen). Briefly, genes encoding the protein of interest were cloned into pFastBac1. Bacmids were obtained from DH10Bac *E. coli* after transformation with the pFastBac1 containing CENP-B ΔC-6xHis-EGFP and used to transfect Sf9 cells (72 hr) to produce baculovirus. Hi-Five insect cells were infected by CENP-B baculovirus (1:10 dilution) for 40 hr, collected, and frozen. The pellets were resuspended in Buffer A. MBP-CENP-B bacteria and CENP-B ΔC-6xHis-EGFP Hi-Five cells were lysed respectively by sonication or cell homogenizer, then centrifuged for 30 min at 45000 g. The clear lysate was loaded onto a 5 ml MBP-Trap (GE Healthcare) for FL MBP-CENP-B and MBP-CENP-B$^{\Delta C}$ or onto a 5 ml His-Trap (GE Healthcare) for CENP-B ΔC-6xHis-EGFP, connected to an Äkta Pure purification system (GE Healthcare). Non-specific proteins were removed by washing the column with buffer A, the proteins of interest were eluted with a linear gradient of Buffer B. For CENP-B ΔC-6xHis-EGFP, an anion exchange chromatography was performed on a Capto Q Impress column (GE Healthcare), the protein of interest was eluted with a linear gradient of Buffer D. A Gel filtration was then performed on the fractions containing eluted proteins using buffer C.

MBP-CENP-B FL and ΔC purification buffers:

- A: 20 mM HEPES pH 7.5, 500 mM NaCl, 1 mM Dithiothreitol (DTT), protease inhibitor (Roche), 1 mM phenylmethylsulfonyl fluoride (PMSF)
- B: 20 mM HEPES pH 7.5, 500 mM NaCl, 1 mM DTT, 1 mM PMSF, 20 mM Maltose
- C: 20 mM HEPES pH 7.5, 500 mM NaCl, 1 mM DTT, 1 mM PMSF, 5% glycerol

CENP-B ΔC-6xHis-EGFP purification buffers:

- A: 20 mM HEPES pH 7.5, 200 mM NaCl, 2 mM ß-Mercaptoethanol, 1 mM PMSF, protease inhibitor (Roche), 10 mM imidazole pH 7.5
- B: 20 mM HEPES pH 7.5, 200 mM NaCl, 2 mM ß-Mercaptoethanol, 1 mM PMSF, 500 mM imidazole pH 7.5

- C: 20 mM HEPES pH 7,5, 200 mM NaCl, 1 mM DTT, 1 mM PMSF, 5% glycerol
- D: 20 mM HEPES pH 7.5, 1 M NaCl, 2 mM ß-Mercaptoethanol, 1 mM PMSF

Stoichiometry of the complexes

The absolute molar masses of proteins were determined using size-exclusion chromatography combined with multi-angle light scattering (SEC-MALS). Protein samples (80 µl at approximately 2 mg/ml) were loaded onto a Superdex 10/300 Increase column (GE Healthcare) developed in 20 mM HEPES pH 7.5, 150 mM KCl, 2 mM $MgCl_2$, 1 mM DTT and 5 mM NaN_3 at 0.5 ml.min^{-1} using a Dionex UltiMate 3000 HPLC system. The column output was fed into a DAWN HELEOS II MALS detector (Wyatt Technology). Data were collected and analysed using Astra X software (Wyatt Technology). Molecular masses were calculated across eluted protein peaks.

Binding assays

Microscale thermophoresis (MST) measurements were performed on a Monolith NT.115 (Nanotemper Technologies) using 57 bp DNA labelled with Alexa Fluor 546 at 5' of one strand. Two-fold dilution series (16 in total) of FL MBP-CENP-B or MBP-CENP-B$^{\Delta C}$ were performed in the interaction buffer (20 mM Hepes pH 7.5, 150 mM NaCl, 2 mM $MgCl_2$, 1 mM DTT and 0.05% (v/v) Tween 20). The DNA was kept at a constant concentration of 100 nM. The samples were loaded into premium capillaries (Nanotemper Technologies). Initial fluorescence at wavelength of 600 nm was measured with excitation at 80% of green LED (550 nm wavelength), at 25°C. CENP-B concentration-dependent fluorescence quenching was observed. To rule out any material loss of fluorescent DNA caused by nonspecific adsorption to tube walls or due to aggregation, an SDS-denaturation test was performed following the manufacturer's instructions. Then the affinity was quantified then by analyzing the change in initial fluorescence of labelled DNA as a function of the concentration of the titrated protein using the MO.AffinityAnalysis v2.1.5 software provided by the manufacturer.

For EMSA, centromeric DNA was incubated at a constant concentration of 8.6 nM together with various concentrations of recombinant CENP-B constructs, in 20 mM HEPES pH 7.5, 100 mM KCl, 2 mM $MgCl_2$, 1 mM DTT. Following a 30 min. incubation at RT, samples were supplemented with 25 % Glycerol prior to the migration on an 0.7% agarose gel, in the electrophoresis buffer (0.25X TBE, 5% Glycerol), during 120 min. at 150V. The gel was stained in an electrophoresis buffer bath supplemented with 1X SYBR safe (Invitrogen life technologies, Carlsbad, CA, USA) during 30 min. at RT, and washed 3 times with the same volume of electrophoresis buffer prior to exposure.

Optical tweezers

CenX and CenY DNA were amplified by PCR using a biotinylated forward primer and a 3-dig reverse primer. The PCR products were then purified from an agarose gel using Macherey Nagel gel and PCR clean-up kit. pKYB1 DNA with biotin-labels at both ends was produced as described earlier (Hoekstra et al., 2017).

All optical tweezers experiments were performed on a commercial instrument (C-trap, Lumicks). 20 ng of centromeric DNA was incubated with 1 µl of anti-digoxigenin coated beads (AD-beads) of 1,03 µm diameter provided by Lumicks. This mix was then diluted in 1000 µl Interaction Buffer (IB: 20 mM Hepes pH 7,5, 150 mM KCl, 4 mM $MgCl_2$, 10 mM DTT) and injected on channel 1. The channel 2 was filled with IB with or without 500 nM of FL-CENP-B or CENP-B$^{\Delta C}$, supplemented with glucose oxydase and glucose. Channel 3 was filled with streptavidin coated beads (SA-beads) of 1,76 µm provided by Lumicks. The pKYB1 DNA was stretched between two SA beads of 1,76 µm. To that end, the beads were flushed into channel 1. pKYB1 DNA (25 ng) was diluted with IB buffer (1000 µL) and directly injected into channel 2, where it was caught from the flow. Stretch experiments were then performed in channel 3, which contained only IB buffer.

For all experiments, the trapping LASER was used at 100% relative power with an overall power of 30%.

Analysis of optical tweezers data

All optical tweezers data was analyzed using python (Hunter, 2007).

To determine the contour length of the stretched DNA the extensible worm-like chain (eWLC) model by Gross et al. (2011) and Odijk (1995) was used

$$d = L_c\left(\left(1-\frac{1}{2}\sqrt{\frac{k_BT}{F\cdot L_p}}\right)+\frac{F}{S}\right), \qquad \text{(Equation 1)}$$

with the distance d, the contour length L_C, Boltzmann's constant k_B, the temperature T, the force F, the persistence length L_P, and the stretch modulus S. Therefore, Equation 1 was solved for the force F and fitted to the initial segment of every force-distance curve (up to 30 pN or the first step in the curve). To increase the robustness of the fit the stretch modulus was fixed to S = 1500 pN (Gross et al., 2011).

To analyze force-clamp experiments the measured bead-bead distance d was converted into the apparent contour length L_C using Equation 1 and the mean persistence length determined for Centromere X DNA in the absence of CENP-B L_P = 26 nm. To extract the DNA length at a given force F_{set} all data points with a force F in the interval [F_{set}-0.25 pN, F_{set}+0.25 pN] were considered.

Average stretch curves were calculated by interpolating the measured distances to an evenly spaced force-axis with 1 pN spacing. To normalize the average distances were divided by the estimated contour lengths (1.02 μm for CenX, 1.00 μm for CenY, 2.75 μm for pKYB1).

AFM sample preparation

All samples were prepared in AFM buffer containing: 50 mM Hepes, 100 mM KCl, 2 mM $MgCl_2$, 1 mM DTT, pH 7.5. A control sample, without proteins, was prepared by mixing 1 μl of (10 ng/μl) centromeric DNA (diluted in AFM buffer) with 8 μl of AFM buffer and 1 μl of (2 mM) spermidine (diluted also in AFM buffer). The whole mix (10 μl) was deposited on freshly cleaved mica for 2 min at room temperature. The mica was then rinsed with 1 ml of ultrapure water and dried under a gentle flow of compressed filtered air.

For protein-DNA constructs first the CENP-B was diluted in the AFM buffer to the desired concentration in AFM buffer (CENP-B concentration 4,5 – 18 nM). Protein-DNA samples were prepared by mixing 1 μl of protein dilution with 7 μl AFM Buffer and 1 μl of (2 mM) spermidine and only after mixing the proteins solution well, 1 μl of template DNA (10 ng/μl) was added to the solution. The whole mix (10 μl) was first pre-incubated for 2 min at room temperature, then deposited on freshly cleaved mica for 2 min at room temperature. Finally, the mica was rinsed with 1 ml of ultrapure water and dried under a gentle flow of compressed filtered air.

AFM imaging

Images were collected using a Bruker Multimode AFM, equipped with a Nanoscope V controller and Nanoscope version 9.2 software. We used Bruker Peakforce-HIRS-F-A probes with nominal stiffness 0.35 N/m and tip radius 1 nm. The imaging was done in PeakForce Tapping mode, with a 4 kHz oscillation frequency, and a peak force setpoint value less than 100 pN. All recorded AFM images consist of 5x5 μm scan areas with 1024x1024 pixels recorded with typical scan frequencies below 1Hz.

Each protein-DNA binding experiment was performed at least in biological duplicate. AFM images were obtained at several separate locations across the mica surface to ensure a high degree of reproducibility and were used for statistical analysis of protein-DNA complexes. Only DNA molecules that were completely visible in AFM image were considered for statistical analysis. AFM images were simply flattened using the Gwyddion software (Version 2.55) without further image processing (Nečas and Klapetek, 2012).

Quantification and statistical analysis of AFM images

Statistical properties such as the contour length and the radius of gyration of DNA molecules (typically between 80–100 individual molecules) were analyzed using "DNA Trace" software (Mikhaylov et al., 2013). We measured the effective persistence length *lp* of control and protein bound DNA molecules by using the bond correlation function for polymers in two-dimensions,

$$<\cos\theta(s)> = e^{-s/2lp}$$

where θ is the angle between the tangent vectors to the chain at two points separated by the distance *s*, and *lp*, the persistence length (Rubinstein and Colby, 2003). Since the persistence length of DNA is obtained by treating the molecules assuming their mechanical properties are uniform (and thus ignoring locally bound proteins), we denote it as the effective persistence length. This analysis is a proven method to extract the persistence length for polymer molecules that orient themselves in an unconstrained way at a surface in equilibrium (Witz et al., 2008). We apply the approach here as well to nucleoprotein filaments that contain small loops, and hence denote *lp* as the effective persistence length.

The square of the radius of gyration of a polymer was defined as the sum of the squares of the distances from all monomers to the center of mass, divided by the total number of monomers:

$$R^2 = \frac{1}{N}\sum_{i}^{N} r_i^2$$

where N is the number of monomers in the chain, and r_i is the distance from the ith monomer to the center of mass (Grosberg et al., 1995).

Nanopore experiments

We used TEM-drilled 7 and 20 nm diameter SiN nanopores (Goeppert) for the experiments. The SiN membrane containing the nanopore was loaded in a PEEK (Polyether ether ketone) flow cell. The DNA samples were diluted in LiCl solution (final concentration 1M) to a final concentration of 2-2.5ng/μl before being introduced to the cis side (-ve) of the nanopore. A positive bias (100-200mV) was applied to drive the DNA sample electrophoretically through the pore.

We used Ag/AgCl electrodes and an Axopatch 200B amplifier (Molecular Devices) for current detection. The traces were recorded at 100 kHz and further filtered at 50kHz with the Transanalyzer Matlab package (Plesa and Dekker, 2015). We set a current blockade threshold at three times higher than the average current blockade of a single DNA helix event (threshold = $3xI_{DNA}$) and quantified the percentage of events above this threshold as compared to the control. All experiments were repeated in duplicates.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical details, which include the value of n (where n represents the number of centromeres measured) can be found in the figure legends and/or in Table S1. All statistical tests were performed using GraphPad Prism 9.0 for Mac (GraphPad Software, San Diego, California, USA, www.graphpad.com).