Base-pair resolution analysis of the effect of supercoiling on DNA flexibility and recognition

Alice L. B. Pyne[1,2†*], Agnes Noy[3,†*], Kavit Main[2], Victor Velasco-Berrelleza[3], Michael M. Piperakis[4], Lesley A. Mitchenall[4], Fiorella M. Cugliandolo[4], Joseph G. Beton[2,5], Clare E.M. Stevenson[4], Bart W. Hoogenboom[2,6], Andrew D. Bates[7], Anthony Maxwell[4], Sarah A. Harris[8,9*]

[1] Department of Materials Science and Engineering, University of Sheffield, Sheffield S1 3JD, UK

[2] London Centre for Nanotechnology, University College London, London, WC1H 0AH, UK

[3] Department of Physics, Biological Physical Sciences Institute, University of York, York, UK

[4] Department of Biological Chemistry, John Innes Centre, Norwich NR4 7UH, UK

[5] Department of Crystallography, Institute of Structural and Molecular Biology, Birkbeck, University of London, London WC1E 7HX, UK

[6] Department of Physics and Astronomy, University College London, London, WC1E 6BT, UK

[7] Institute of Integrative Biology, University of Liverpool, Liverpool L69 7ZB, UK

[8] School of Physics and Astronomy, University of Leeds, Leeds, LS2 9JT, UK

[9] Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds LS2 9JT, UK

† authors contributed equally

* Email: a.l.pyne@sheffield.ac.uk, agnes.noy@york.ac.uk, s.a.harris@leeds.ac.uk

Present addresses:

Michael Piperakis, CU Coventry, 8 Mile Lane, Coventry, CV1 2TU, UK.

Fiorella Cugliandolo, Division of Immunology, Department of Pathology, University of Cambridge, Cambridge, CB2 1QP, UK

Abstract

In the cell, DNA is arranged into highly-organised and topologically-constrained (supercoiled) structures. It remains unclear how this supercoiling affects the detailed double-helical structure of DNA, largely because of limitations in spatial resolution of the available biophysical tools. Here, we overcome these limitations, by a combination of atomic force microscopy (AFM) and atomistic molecular dynamics (MD) simulations, to resolve structures of negatively-supercoiled DNA minicircles at base-pair resolution. We observe that negative superhelical stress induces local variation in the canonical B-form DNA structure by introducing kinks and defects that affect global minicircle structure and flexibility. We probe how these local and global conformational changes affect DNA interactions through the binding of triplex-forming oligonucleotides to DNA minicircles. We show that the energetics of triplex formation is governed by a delicate balance between electrostatics and bonding interactions. Our results provide mechanistic insight into how DNA supercoiling can affect molecular recognition.

Introduction

Native genomic DNA in prokaryotes is maintained under negative-superhelical stress, a state of torsion that tends to unwind the DNA double helix[1–3]. The DNA response to this torsional stress is described as supercoiling, in which a change in linking number (ΔLk) of the DNA is partitioned into twist (Tw) and writhe (Wr)[1–4]. In prokaryotes, genomic DNA is globally negatively supercoiled, with superhelical density $\sigma \approx -0.06$[5]. This supercoiling operates synergistically with nuclear-associated proteins to regulate bacterial gene expression[6]. In eukaryotes, supercoiling is implicated in the regulation of oncogenes such as c-Myc[7]. It plays a fundamental role in the formation and stability of looped DNA structures[8] and DNA R-loops[9], and influences the placement of RNA guide sequences by the CRISPR-Cas9 gene editing toolkit[10]. The supercoiling-induced structural changes that modulate these DNA functions present a challenge for traditional structural methods that can provide atomistic resolution, *i.e.*, X-ray crystallography[11] and nuclear magnetic resonance (NMR)[12], because of the diverse conformational landscape of supercoiled DNA[13].

As part of its role in regulating transcription, replication, and chromosomal segregation[14], supercoiling has been proposed to play a role in the specificity of DNA-binding ligands, including major-groove binders such as triplex-forming oligonucleotides (TFOs)[15,16]. TFOs target specific DNA sequences, forming a triplex of the single-stranded TFO and the target duplex DNA[17]. The target specificity of TFOs combined with their ability to suppress gene expression has driven their development as anti-cancer agents.

Here we combine high-resolution atomic force microscopy (AFM) with molecular dynamics (MD) simulations to reveal how supercoiling affects global and local DNA conformation, structure and dynamics in DNA minicircles of length 250-340 bp. These minicircles are small enough to be simulated at the atomistic level by MD[13,18] and to be visualized at high (double-helix) resolution by AFM experiments in solution[19–21]. Minicircles are also representative of looped DNA at plectoneme tips[22] and small extrachromosomal circular DNAs (eccDNAs), which have tissue-specific populations and sequence profiles in human cells[23–25]. The DNA minicircles in this study incorporate a TFO-binding sequence, to assess how the interplay of electrostatic and base-stacking energies determines the formation of triplex structures in supercoiled DNA.

Results

High-Resolution AFM and MD reveal conformational diversity in supercoiled DNA minicircles: Figure 1 shows the structure of negatively-supercoiled DNA minicircles as viewed by high-resolution AFM and simulated by atomistic MD. High-resolution AFM images recorded in aqueous solution show DNA minicircles, isolated with native levels of supercoiling, in a range of conformations with sufficient resolution to resolve the two oligonucleotide strands of the double helix. For the 251 bp minicircle, this allowed

determination of the linking number, Lk = 24 ± 1 from direct measurements of twist (24 ± 1 turns) and writhe (≤ 1). The measured twist corresponds to a helical repeat of 10.5 ± 0.5 bp, consistent with canonical B-form DNA[1]. For each conformation of the surface-bound minicircles found by AFM (Fig. 1a-d), it was possible to find MD-generated conformers with a close resemblance in global structure (Fig. 1e). The high planarity of the minicircles, determined in the simulations as less than 15% on average (Supp. Fig. 1), is advantageous to structural determination by AFM, because distortions resulting from surface immobilisation are minimal for planar molecules. Atomistic models of supercoiled DNA minicircles have been shown to be consistent with cryo-ET density maps[13], which provide sufficient resolution to capture the overall shape of the minicircles, but not their helical structure. The variation in structures observed in Fig 1a-e is attributed to thermal fluctuations within supercoiled DNA, with time-resolved AFM (Fig. 1f) demonstrating that dynamic behaviour can occur in these molecules on the order of minutes, even when tethered to a surface, also implying that the structural perturbation caused by immobilisation is small. Similar dynamics were observed in MD simulations of the 339 minicircle (ΔLk = -1) in a continuum representation of the solvent, albeit at a much faster (picosecond) rate (Fig. 1g and Supp. Videos 1, 2). Experimental measurements have shown that adsorption to a surface for AFM slows dynamics[26,27] and *in silico*, the absence of friction with water molecules accelerates conformational dynamics[13,28].

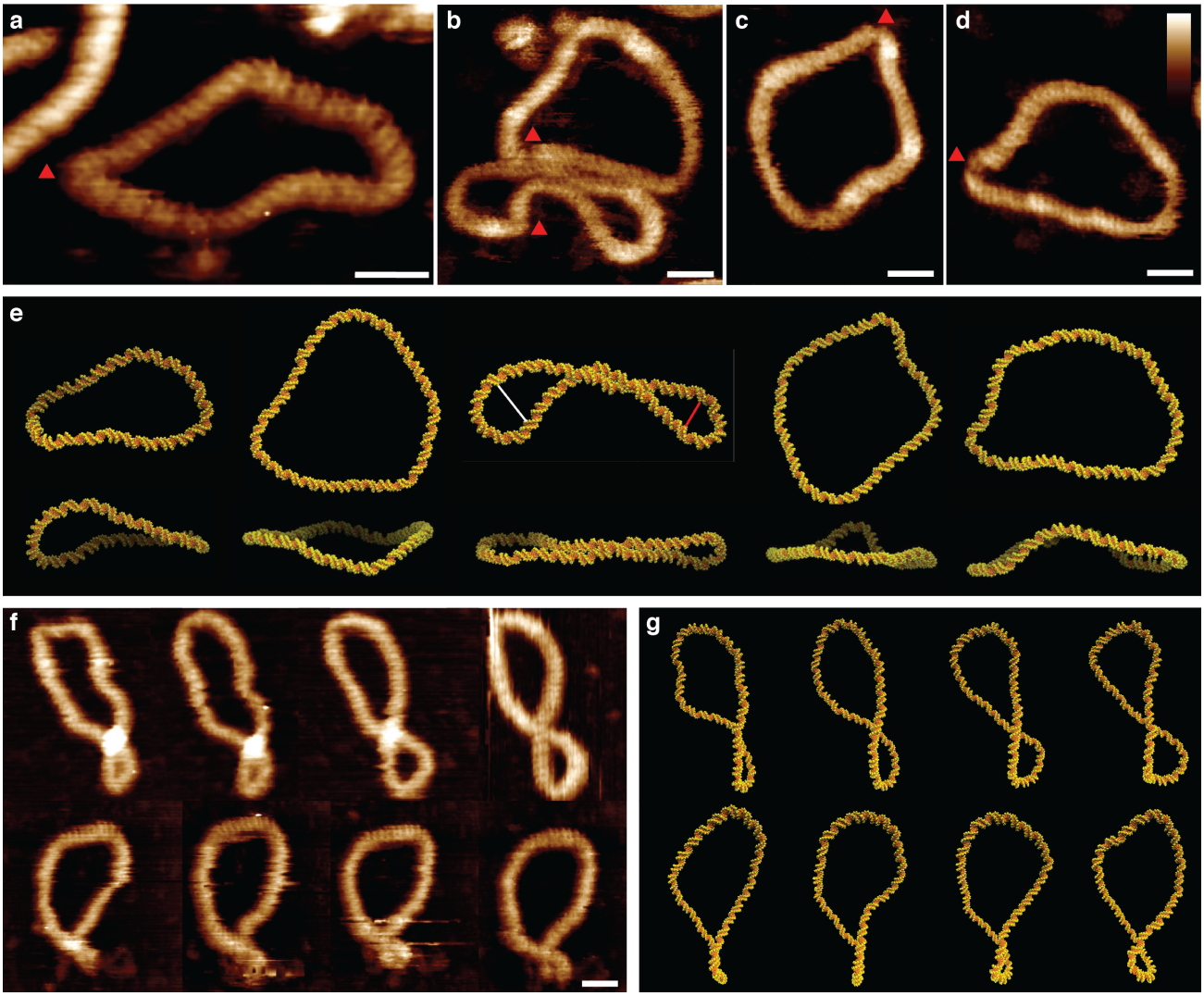


Figure 1: Structural and dynamic diversity in supercoiled DNA minicircles. **a-d,** High resolution AFM images of natively-supercoiled (σ = 0.03-0.06) DNA minicircles of 251 (**a**) and 339 (**b**, **c**, **d**) bp showing their helical structure and disruptions of canonical B-form DNA (marked by red triangles), where the angle of the helix changes rapidly, or where the DNA appears thinner or disrupted. **e**, MD snapshots of minicircle conformations for 251 (1st image) and 339 bp corresponding to the minicircles in the AFM images, selected from simulations at ΔLk 0, -1 and -2. Top and side views (top and bottom row, respectively) show the degree of planarity of the depicted structures. White and red lines indicate plectonemic loops of 9 and 6.5 nm width, respectively (see methods). **f**, Time-lapse AFM measurements of a natively-supercoiled 339 bp DNA minicircle, recorded at 3 min/frame. **g**, Chronological snapshots from simulations of 500 ps duration for a 339 bp minicircle with ΔLk = -1 (see Supp. Videos 1 and 2). Scale bars (inset) 10 nm, height scale (inset d): 2.5 nm for all AFM images.

Negative supercoiling induces defects in DNA minicircles: AFM not only provides resolution sufficient to observe the DNA helical repeat, but most critically achieves this without the need for ensemble averaging. This uniquely permits us to observe heterogeneous structural perturbations, e.g. individual DNA defects,

that occur due to superhelical stress imposed on the minicircle. By combining AFM and *in silico* measurements of DNA minicircle topoisomers with increasing levels of supercoiling (Fig. 2) we were able to observe the effect of negative supercoiling on the structure and mechanics of DNA with Ångström resolution. We observed no defects in the structure of the relaxed topoisomer, which maintains a B-form structure throughout the molecule. However, in negatively-supercoiled DNA, defects were observed both by AFM (Fig. 1 a-d, red triangles) and atomistic MD simulations (Fig. 2a, red triangles). We observed the onset of defects in negatively-supercoiled minicircles of $\Delta Lk = -1$ onwards ($\sigma \approx -0.03$); across all *in silico* topoisomers, seven out of the ten defects observed are denaturation bubbles, where two or more base pairs are flipped out of the duplex (Fig. 2a insets and Supp. Fig. 2). This results in flexible hinges that can accommodate a 180^0 turn within a single helical turn, radically altering the range of conformations the DNA can adopt. We also observed type I kinks in topoisomers -1 and -3 (in which a single base pair presents a strong bend, breaking hydrogen bonds and stacking), and a type II kink within topoisomer -3 (in which hydrogen bonds of two consecutive base pairs are broken and bases are stacked on their 5´ neighbours) (Fig. 2a insets)[29].

Direct comparison of the level of negative supercoiling required to induce the onset of structural transitions, including denaturation of unbent DNA (typically taken to be around $\sigma \approx -0.04$[30]) with that of DNA minicircles is not straightforward, because the DNA supercoiling response is so exquisitely sequence-dependent. For longer sequences, the statistical likelihood that a sequence will contain an element that undergoes a specific stress-induced structural transition (e.g. Z-DNA formation or cruciform extrusion) is larger[31], and these element suppress defect formation by absorbing superhelical stress[32]. Our minicircle sequences do not contain any such supercoiling responsive sequences. The 1 or 2 bp size defects we observe in minicircles are smaller than those that have been probed in 2-5 kbp negatively-supercoiled plasmids (> 30 bp)[31]. However, coarse-grained simulations of 600 bp supercoiled linear DNA show the formation of small defects (2-3 bp) at plectonemic loops, with larger bubbles (up to 20 bp) observed when plectoneme formation is prohibited by an applied force[29]. We deduce from these observations that DNA bending promotes and localises supercoiling-induced defect formation. Bent DNA structures are ubiquitous in the genome; as well as forming ends of plectonemes[33], bent DNA is vital to a number of recognition processes, including transcription regulation via DNA looping[34], and DNA damage detection[35,36].

Estimate of critical bend angle associated with defect formation: We determined the critical bending angle required to form a defect through curvature analysis for all *in silico* topoisomers (Fig. 2a, b) and for natively-supercoiled DNA minicircles by high-resolution AFM (Fig. 2c). Kinks were observed by AFM as discontinuities in the helical repeat of DNA where the angle of the helix changes rapidly, or where the DNA appears thinner or disrupted (Fig. 2c). Defects in the MD were classified as disruptions to base stacking and complementary base pairing (Fig. 2a, insets). Figure 2d shows DNA minicircle bend angles classified as either B-form (black crosses) or defective DNA (red triangles), both for AFM (1st column) and MD (all

other columns). We deduce that canonical B-form DNA can sustain an angle of up to ~75° through regions of high bending stress (critical angles of 76° and 74° for AFM and MD respectively – Supp. Fig. 2) without disruption to either base stacking or hydrogen bonding. For defective DNA, an average bend angle of 106 ± 15° was measured for AFM and 120 ± 32° for MD, almost double the bend angle measured for canonical DNA, of 69 ± 5° for AFM and 57± 9° (mean ± standard deviation). This maximum bend angle of 75°, implies that for a DNA bend (such as a plectoneme), to remain free of defects the loop must be more than 7-10 nm wide, which requires approximately 55 bp or five helical turns, showing remarkable similarity with coarse grained simulations[29].

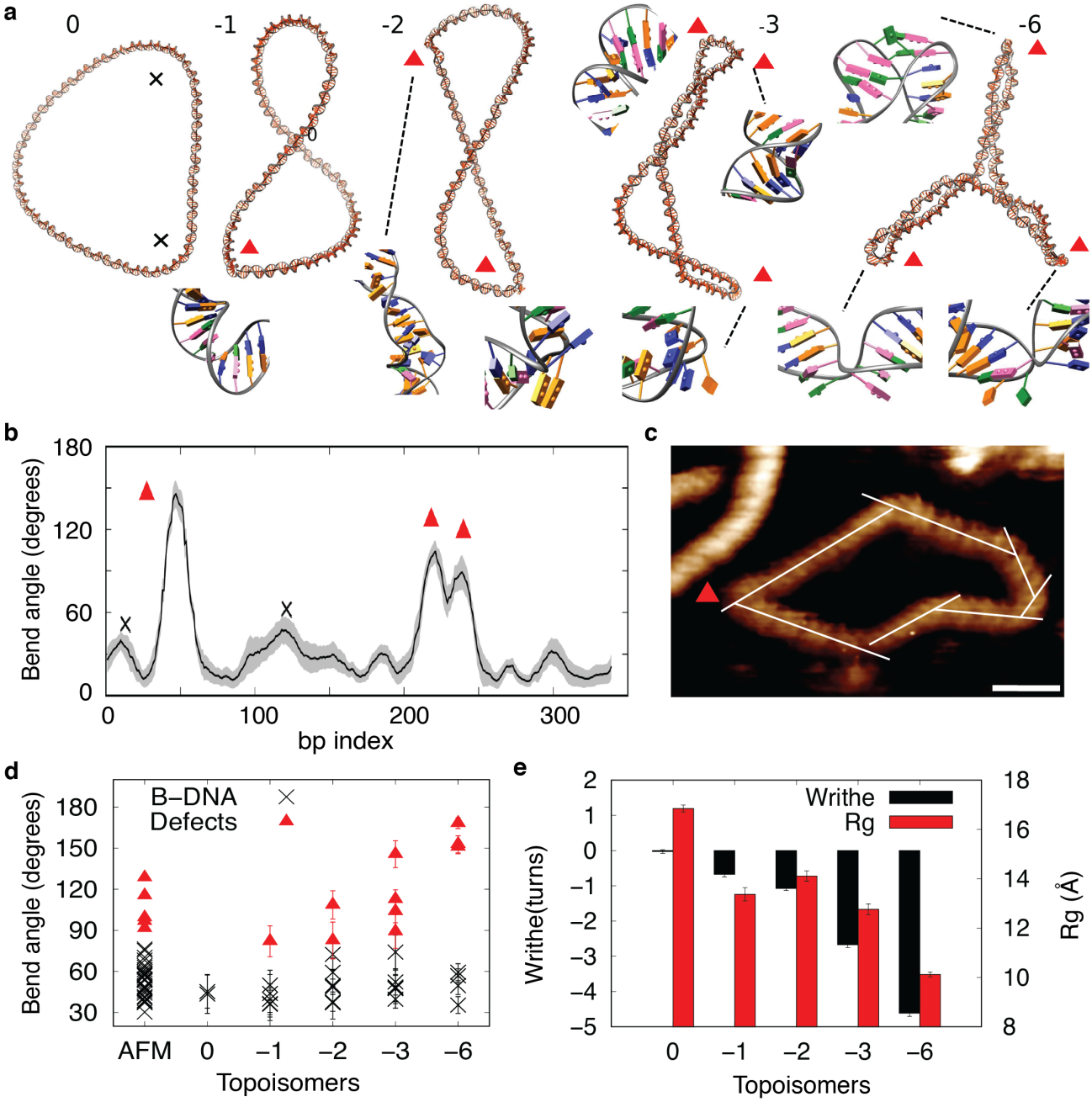


Figure 2: Supercoiling induces defect formation in 339 bp DNA minicircles, while increasing writhe and compaction. **a**, MD average structures showing increased defect formation at higher supercoiling. **b,** Bending calculation obtained by SerraLINE program using WrLINE profile from the -3 topoisomer trajectory, where bend angles are calculated as a directional change in tangent vectors separated by 16 bp (additional bending profiles in Supp. Fig. 2). All peaks higher than 35^0 are classified as B-DNA bends (black cross) or defects (red triangles) depending on whether canonical non-bonded interactions were broken. **c**, Determination of bending angles in natively-supercoiled DNA by high-resolution AFM (white lines) **d** Bent-DNA analysis of DNA minicircles by high-resolution AFM (natively-supercoiled, 1st column), and MD simulations (topoisomers 0 to -6, **a**) shows a $\approx 75^0$ cut-off between B-DNA (black crosses) and defects (red triangles), with an increasing of the latter with supercoiling. **e**, Radius of gyration (Rg) and writhe from MD simulations. Grey shading (**b**) and error bars (**d**,**e**) correspond to standard deviations. AFM scale bar (**c**) 10 nm, height scale (**c**) 2.5 nm.

Global compaction in DNA structure correlates with the formation of defects: To probe how the supercoiling-induced changes in DNA structure vary with the global conformation of DNA minicircles, we generated a range of relaxed and negatively-supercoiled topoisomers experimentally (Fig. 3a,b) for comparison with those generated *in silico* (Fig. 2). For each topoisomer we quantified the degree of molecular compaction observed by AFM (Fig. 3a) and determined the supercoiling as an average of all bands observed by gel electrophoresis (Fig. 3b). Aspect ratios were calculated for individual minicircles within images (Fig. 3a). While relaxed DNA minicircles appear predominantly as open rings, with high aspect ratio, increasing superhelical density increases the global compaction generating a range of heterogeneous structures containing defects (Fig. 3c,d). This global compaction from relaxed to maximally supercoiled structures is accompanied by a decrease in the aspect ratio of 35% by AFM (Fig. 3e) and 40% by MD (Fig. 2e).

As expected, as ΔLk decreases from 0 to -1, ($\sigma \approx 0$ to -0.03) the DNA writhes and compacts. However, further negative-supercoiling of the helix to $\Delta Lk \approx -2$ ($\sigma \approx -0.06$) results in a counterintuitive decrease in compaction (Fig. 3e). This correlates with a smaller electrophoretic shift for -1 to -2 than for the other topoisomer transitions (Fig. 3b) and a smaller change in writhe in the MD simulations than for other transitions (ΔWr = -0.4 turns and -1.7 turns for the -1 to -2 and -2 to -3 transitions respectively). This anomalous behaviour correlates with the onset of defects observed by both AFM and MD. Defects relieve torsional stress and allow the DNA to partially relax, resulting in an increased number of open conformations (Fig. 3c). Comparing the writhe of a defect containing (-1.1 ± 0.1 turns) and defect free simulation (-1.7 ± 0.1 turns) of the $\Delta Lk = -2$ topoisomer shows that defects cause a reduction in writhe of 0.6 turns (conformers shown in Supp. Fig. 2). When further negative supercoiling is introduced, the DNA becomes increasingly writhed and compacted, as the superhelical stress can no longer be dissipated purely through the formation of defects.

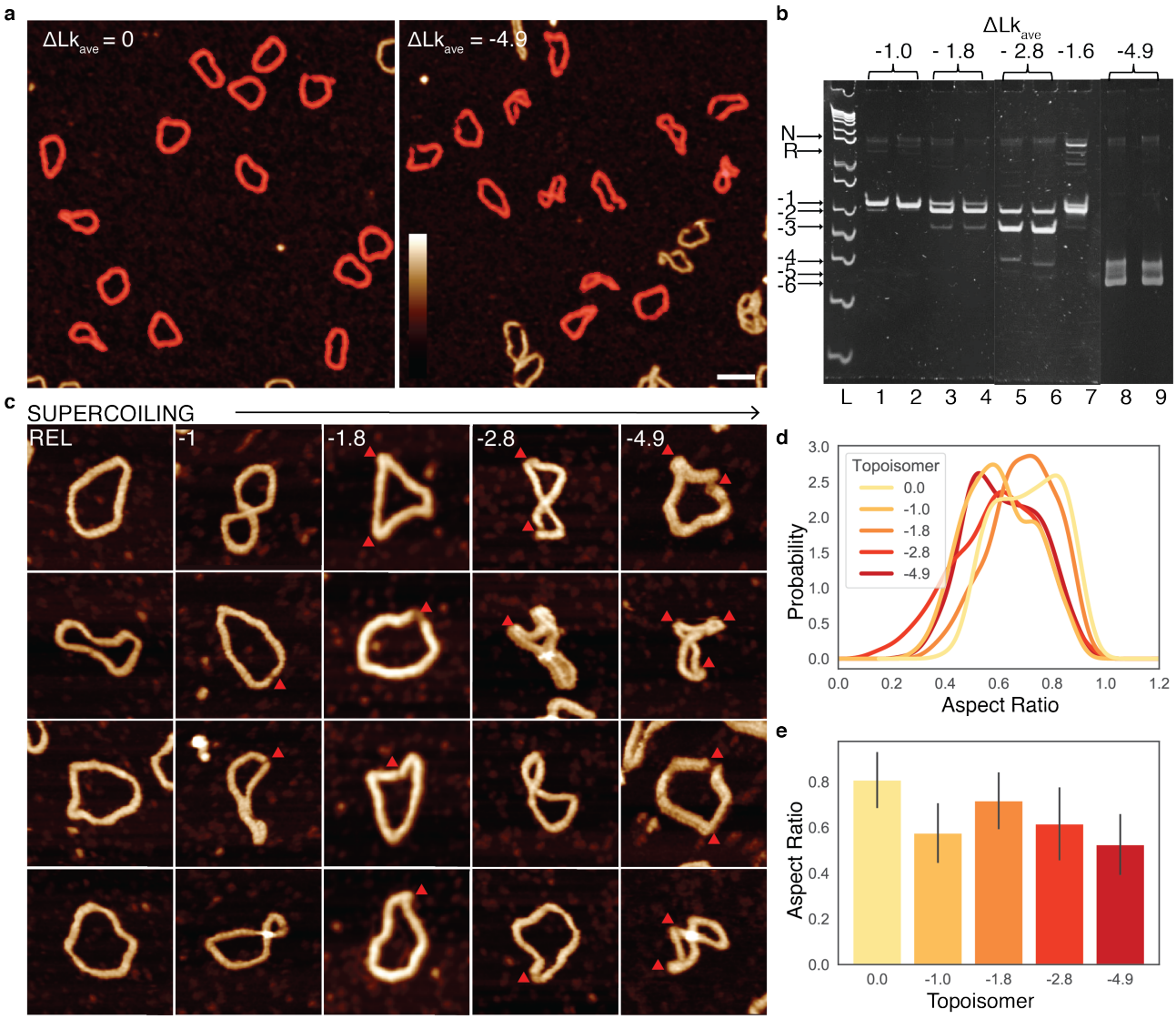


Figure 3: Negative supercoiling induces global compaction of DNA minicircles, with a conformational change observed at physiological levels of supercoiling. **a,** AFM images of DNA minicircle populations show increased writhe and compaction at increased negative superhelical density. Images are processed to obtain individual minicircles (red) for analysis[37]. Height scale (inset) 4 nm, scale bar 50 nm. **b**, 5% TAC acrylamide gel of negatively-supercoiled topoisomers of 339 bp (ΔLk from -1 to -4.9) generated by addition of increasing amounts of ethidium bromide during the re-ligation reaction. N = nicked minicircle; R = relaxed minicircle. **c,** Representative images of 339 bp minicircles for a range of superhelical densities showing increased levels of compaction and defects (observed as regions of high bending angle, or discontinuities in DNA structure, marked by red triangles) for highly-supercoiled minicircles. Height scale (inset, **a**) 4 nm, all images are 80 nm wide. **d,** The relationship between minicircle aspect ratio and supercoiling as a Kernel Density Estimate (KDE) plot of probability distribution for each topoisomer (N = 1375). **e**, Mode aspect ratio for each minicircle topoisomer, as calculated from the maximum of each KDE in **d**, error bars correspond to standard deviation.

Supercoiling-induced conformational variability affects binding of triplex forming oligonucleotides: The effect of supercoiling-induced structural variability on DNA binding interactions was investigated through the site-specific binding of a triplex forming oligonucleotide (TFO) to supercoiled DNA minicircles. The formation of triplex DNA occurs via Hoogsteen base-pairing between the $(CT)_{16}$ triplex-forming oligonucleotide (TFO) and the double-stranded minicircle sequence $(GA)_{12}$[38]. By AFM, we observe triplex formation as small, sub-nanometre protrusions from natively-supercoiled DNA minicircles (Fig. 4a). This was verified by AFM measurements on linearised DNA minicircles (Supp. Fig. 3). Strikingly, SPR experiments showed that the superhelical density of the minicircles has a minimal effect on the affinity of triplex formation, with binding constants (K_d) of the order of 10 pM across superhelical densities ($\Delta Lk = 0$ to -4.9, Supp. Fig. 4 and Supp. Table 1).

To understand the unexpected lack of sensitivity of triplex binding to supercoiling in minicircles, we determined the local and global energetic contributions associated with the binding of the TFO *in silico* for a range of supercoiled topoisomers (Fig. 4b). As expected, the relative contributions of the local electrostatic (Fig. 4c) and the hydrogen bonding and stacking (Fig. 4d) interactions vary with superhelical density. As the DNA minicircles are compacted by superhelical stress, the electrostatic penalty for triplex binding increases (Fig. 4c), due to the increase in local negative charge. Taken alone this would imply that triplex formation is disfavoured by DNA supercoiling; however, the increased electrostatic penalty is offset by the formation of new hydrogen bonds upon triplex formation (Fig. 4d). The new hydrogen bonds (Fig. 4b, inset), preferentially observed in topoisomers of higher negative supercoiling, consist of (i) bifurcated hydrogen bonds between the Watson-Crick (WC) binding pyrimidine strand and the TFO and (ii) C-H::O weak hydrogen bonds between the backbone of the WC binding purine strand and the bases of the TFO (Fig. 4d, blue). In addition, negative supercoiling predisposes DNA to triplex formation, because twist values as low as 30^0 are observed in triplexes, so triplex formation relieves supercoiling by local unwinding. Conversely, positive supercoiling disrupts Hoogsteen H-bonds disfavouring triplex interaction (Fig. 4d, yellow).

Local changes in non-bonded interactions with the TFO induce only a minimal perturbation to the writhe, except in the case of the highly compacted $\Delta Lk = -6$ structure, which shows a reduction of half a helical turn in writhe (Fig. 4e) on triplex binding. This balance between the energetic compensation from the competition of unfavourable electrostatics with increased hydrogen bonding implies that triplex formation should only be minimally affected by the supercoiling induced variation in global structure, as is observed by SPR.

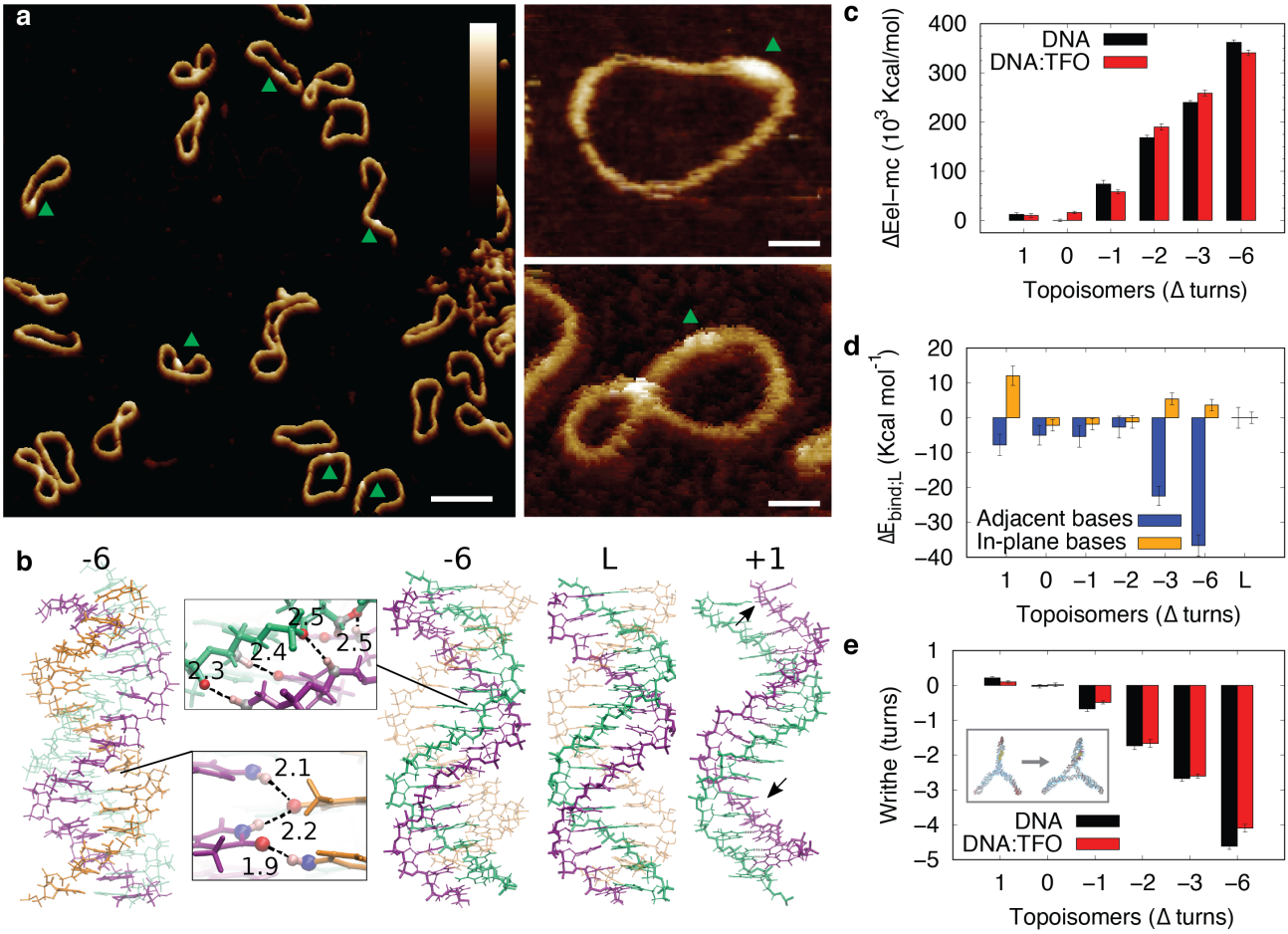


Figure 4: Conformational diversity in supercoiled DNA minicircles contributes to triplex formation. **a,** AFM images showing triplex formation across a range of DNA minicircle conformations. Triplex regions are visible as small, sub-nanometre protrusions from the DNA marked by green triangles. Height scales (scale bar inset): 3 nm, scale bars (single minicircles): 10 nm, scale bars (population): 50 nm. **b,** Representative structures of DNA triplex from -6 and +1 topoisomer simulations compared to linear DNA. Arrows indicate less favourable Hoogsteen hydrogen bonds in positively supercoiled DNA. The WC-pyrimidine strand is erased from ΔLk = +1 image for visualisation purposes. **c**, Averages and standard deviations for electrostatics of the whole minicircle ($\Delta E_{elec;0}$). **d**, Non-bonded interactions for the triplex-binding site ($\Delta E_{bind;L}$), showing the relative contributions from in plane base interactions (e.g. WC and Hoogsteen hydrogen bonds) (yellow bars), compared to interactions between adjacent bases (e.g. bifurcated and backbone hydrogen bonds and stacking energies) (blue bars). **e**, Minicircle writhe for topoisomers with and without TFO bound. Inset shows a half helical turn reduction in writhe on triplex binding for the ΔLk = -6 topoisomer.

Conclusion

Using a combination of high-resolution AFM and atomistic MD simulations, we describe the structure, dynamics and recognition of negatively-supercoiled minicircle DNA, with double-helical resolution. We quantify the critical bend angle for canonical B-form DNA under superhelical stress as 75^0, implying that a

DNA loop must be formed of at least 5 helical turns to be free of defects. These defects dominate DNA mechanics by contributing to the flexibility and conformational diversity of supercoiled DNA.

We observe that superhelical stress globally compacts DNA, resulting in a decreased aspect ratio, and radius of gyration. However, at superhelical densities close to the that of genomic DNA, we see an unexpected reduction in compaction. We attribute this reduction to the onset of supercoiling-induced kinks and defects, through which torsional stress can be dissipated. Beyond this point the trend to compaction continues, as the defects generated are not sufficient to absorb increased superhelical stress.

The conformational diversity of supercoiled DNA allows for structural perturbations that can accommodate the binding of external substrates as exemplified the formation of triplex DNA. The supercoiling dependence of triplex formation is governed by a balance of two competing energetic interactions. An increased electrostatic penalty is incurred in negatively-supercoiled DNA due to supercoiling-induced compaction, whilst additional hydrogen bonds are facilitated by DNA undertwisting. This balance in the energetics facilitates triplex formation across a range of superhelical densities.

We show remarkable synergy between atomistic simulations and experimental data; despite caveats in both, including a requirement for surface binding, and the use of empirical classical forcefields with sampling limitations imposed by finite computational resources. However, in combination these two biophysical tools enable us to determine the effect of supercoiling on local and global DNA structure and its wider influence on dynamics and recognition. Though we note that the bending stress for DNA minicircles is much higher than for longer DNA, such as plasmids, or eukaryotic topologically associated domains (TADs), the significant perturbation on the mechanics of the DNA enforced by the bending energy in minicircles of this size has particular relevance for the structure of tightly constrained DNA for example at plectoneme ends[22,39], short DNA loops and in small eccDNAs. We believe that this data, taken in conjunction with studies of longer DNAs under superhelical stress, will provide a more complete study of DNA structure under stress and can be used to inform future studies on DNA nanotechnology, plectoneme[39] and topology prediction[22]. As well as improving our fundamental understanding of DNA mechanics, our findings have applications in bioengineering, given the proposed therapeutic potential of small circular DNAs and TFOs[40] and the required optimisation of DNA for diagnostics[41] and therapeutics[42].

Methods

Generation and purification of small DNA circles

Small DNA circles (minicircles) of 339 and 251 bp were prepared using bacteriophage λ-Int site-specific recombination in vivo, based on a method previously described with some minor modifications[43]. In each case, a 16 bp triplex-binding site (for the triplex-forming oligo TFO1R: 5´[Bt] –CTC TCT CTC TCT CTC T (where Bt indicates biotin), the reverse of the sequence described previously[16].

Plasmids containing the original minicircle sequences were provided by Lynn Zechiedrich (Baylor College, Houston, Texas). For the 251 bp circles we experienced low yields for the methods described above so most material was obtained from Twister Biotech (Houston, TX, USA); we also obtained larger quantities of 339 bp circles from this company.

The triplex-forming region were incorporated into the parent plasmids by site-directed mutagenesis using the QuikChange Site-Directed Mutagenesis Kit (Stratagene) following the manufacturer's protocol. Plasmids were transformed into *Escherichia coli* LZ54[43]; the 339 bp minicircles were prepared and isolated using three methods.

On a small scale (2 L), plasmid-containing strains were grown in LB medium, as described previously[43] except that BamHI (which linearises the large circular product that is catenated to the minicircle) was not used to release the minicircle; we found that treatment with BamHI did not increase the yield of the minicircle product. (We presume that the action of DNA topoisomerase IV (topo IV) during cell harvesting was sufficient to achieve this.)

On a larger scale (up to 100 L), 2-L cultures (as described above) were used to inoculate 100 L of modified Terrific Broth in a bioreactor at the Wolfson Fermentation and Bioenergy Laboratory (University of East Anglia, Norwich, UK). The modified Terrific Broth contained 12 g tryptone, 48 g yeast extract, 30 mL glycerol, 0.1 mL antifoam 204 (Sigma-Aldrich), 2.32 g KH_2PO_4 and 12.54 g K_2HPO_4 per litre; ampicillin was added to a final concentration of 100 μg/mL. Cells were grown at 30 °C and the pH maintained at 7.0 during growth by the addition of 5 % (v/v) phosphoric acid when needed. The dissolved oxygen concentration was maintained at >40% by agitation control. Cells were grown to mid-exponential phase (A_{600} = 3.5) at which point Int expression was induced by shifting the cultures to 42 °C for 30 min. Norfloxacin (Sigma-Aldrich) was then added to 30 μg mL^{-1} and the cultures were shifted back to 30 °C. After 1 h, the cells were harvested by centrifugation and the pellet split into 10 batches (180 g per batch); the protocol below describes the procedure carried out for each of the cell pellet batches.

The cell pellet was resuspended in 500 mL of 25 mM Tris·HCl (pH 8.0), 50 mM glucose, 10 mM EDTA, and was incubated at room temperature with 2.5 mg mL^{-1} lysozyme (Sigma-Aldrich, chicken egg white)

for 30 min. The cells were then lysed by addition of 1 L 1% SDS, 0.2 M NaOH for 5 min at room temperature, after which 750 mL of 3 M potassium acetate (pH 4.0), was added. Protein precipitation was allowed to occur for >1 hour at 4°C. Cell debris was removed by centrifugation, and the supernatant was filtered through miracloth under vacuum. Nucleic acid was next precipitated by the addition of isopropanol (0.7 vol) to the filtrate. The resulting harvested pellet was resuspended in 120 mL 10 mM Tris·HCl [pH 8.0], 1 mM EDTA and an equal volume of 5 M LiCl added to precipitate high molecular weight RNA, which was removed by centrifugation. The supernatant was precipitated with ethanol, air dried, resuspended in 150 mL 50 mM MOPS pH 7.0, 5 mM EDTA and then treated with RNase A (Sigma-Aldrich, 50 μg/mL) for 30 min at 37°C, followed by proteinase K (Sigma-Aldrich, 50 μg mL^{-1}) for a further 30 min at the same temperature. Most of the unwanted large circle was removed by PEG precipitation; to the DNA suspension, 150 mL of 10 % PEG-8000, 1.5 M NaCl was added and the resulting mixture was incubated at 4°C for 15 min. The mixture was centrifuged, and the supernatant was treated with 200 mL anion-exchange loading buffer (50 mM MOPS [pH 7.0], 750 mM NaCl, 5 mM EDTA) to reduce the PEG concentration. The DNA minicircles were isolated on QIAGEN-tip 10000 anion-exchange columns following the manufacturer's guidelines. The isolated minicircle was then subjected to Sephacryl S-500 gel filtration to further purify it. Fractions containing minicircle DNA were pooled, and concentrated by isopropanol precipitation, washing the precipitate with ethanol. Purification by gel filtration was repeated a few successive times in order to ensure complete removal of dimeric minicircle. The purified and concentrated minicircle DNA was re-suspended in TE buffer.

DNA minicircle sequences:

AttR sequence is highlighted in blue, and triplex forming sequence in green.

251 bp:

TTTATACTAACTTGAGCGAAACGGGAAGGTAAAAAGACAACAAACTTTCTTGTATACCTTTAAGAGAGAGAGAGAGAGAGACGACTCCTGCGATATCGCCTCGGCTCTGTTACAGGTCACTAATACCATCTAAGTAGTTGATTCATAGTGACTGCATATGTTGTGTTTTACAGTATTATGTAGTCTGTTTTTTATGCAAAATCTAATTTAATATATTGATATTTATATCATTTTACGTTTCTCGTTCAGCTTT

339 bp:

TTTATACTAACTTGAGCGAAACGGGAAGGGTTTTCACCGATATCACCGAAACGCGCGAGGCAGCTGTATGGCGAAATGAAAGAACAAACTTTCTTGTACGCGGTGGTGAGAGAGAGAGAGAGAGATACGACTACTATCAGCCGGAAGCCTATGTACCGAGTTCCGACACTTTCATTGAGAAAGATGCCTCAGCTCTGTTACAGGTCACTAATACCATCTAAGTAGTTGATTCATAGTGACTGCATATGTTGTGTTTTACAGTATTATGTAGTCTGTTTTTTATGCAAAATCTAATTTAATATATTGATATTTATATCATTTTACGTTTCTCGTTCAGCTTT

Preparation and analysis of different topological species of minicircles

To generate negatively-supercoiled species, the 339 bp minicircle was first nicked at a single site using Nb.BbvCI (New England Biolabs) at 37°C. After incubation at 80°C for 20 min to inactivate the endonuclease, nicked DNA was purified and isolated using the QIAGEN miniprep kit. Then, 15 μg of the purified nicked minicircle was incubated with T4 ligase (New England Biolabs) and ligase buffer containing 25 μg mL^{-1} BSA, in the presence of different quantities of ethidium bromide (EtBr) in a total reaction volume of 3 mL, at room temperature overnight. This was followed by successive purification and isolation of pure supercoiled minicircle DNA using both the QIAGEN nucleotide removal and miniprep kits. The average ΔLk (linking number difference) for each species was determined by calculating the weighted average of all closed-circular forms by measuring the intensity of each respective band on a polyacrylamide gel (Fig. 3b) the linking number difference (ΔLk) of each species was assigned by counting bands on gels, as follows: Lane 1: ΔLk_{ave} = -1.0; Lane 2: ΔLk_{ave} = -1.0; Lane 3: ΔLk_{ave} = -1.8; Lane 4: ΔLk_{ave} = -1.8; Lane 5: native supercoiled (ΔLk_{ave} = -1.6); Lane 6: ΔLk_{ave} = -2.8; Lane 7: ΔLk_{ave} = -2.8; Lane 8: ΔLk_{ave} = -4.9; Lane 9: ΔLk_{ave} = -4.9; L = 1 kbp plus ladder (ThermoFisher Scientific).

Linear forms were prepared by digestion with restriction enzyme NdeI (New England Biolabs); relaxed forms were generated either using wheat-germ topoisomerase I (Promega) or by the nicking/ligation procedure described above in the absence of EtBr. DNA samples were analysed by electrophoresis through 5 % polyacrylamide gels (acrylamide/bis = 29:1) in TAC (40 mM Tris·acetate [pH 8.0], 10 mM $CaCl_2$) or TAE (40 mM Tris·acetate [pH 8.0], 1 mM EDTA) at 100 V for ~3 h. Gels were stained with SYBR Gold (Invitrogen) and analysed using a Molecular Dynamics STORM 840 imaging system with quantitation using ImageQuant software.

Plasmid pBR322 was supplied by Inspiralis Ltd (Norwich, UK) and analysed by electrophoresis through 1 % agarose gels in TAE buffer at 80 V for ~2 h. Gels were stained with EtBr and analysed using a Molecular Dynamics STORM 840 imaging system with quantitation using ImageQuant.

S1 nuclease digestions

To determine whether triplex formation between TFO1R and minicircle DNA had occurred, samples were probed with S1 nuclease. To prepare the triplex complex, an excess of TFO1R (2.5 μM) was incubated with the minicircle/plasmid (150 nM) in 100 mM calcium acetate pH 4.8, in a total volume of 20 μL at room temperature for 30 min. (In control experiments reactions were also carried out in TF buffer: 50 mM sodium acetate pH 5.0, 50 mM NaCl, 50 mM $MgCl_2$.) Aliquots (5 μL) were taken and S1 nuclease (0 to 1000 U; Thermo Fisher Scientific) was then added and the incubation continued in S1 nuclease buffer (30 mM sodium acetate pH 4.6, 1 mM zinc acetate, 50 % [v/v] glycerol) at room temperature for 30 min; the total volume of these reactions was 10 μL. The digest was stopped by the addition of 0.25 M EDTA (5 μL)

followed by heat inactivation at 70 °C for 10 minutes; DNA was isolated by extraction with chloroform:isoamyl alcohol.

Atomic Force Microscopy

DNA minicircle sample preparation for AFM imaging:

DNA minicircles were adsorbed onto freshly cleaved mica specimen disks (diameter 6 mm, Agar Scientific, UK) at room temperature, using either Ni^{2+} divalent cations or poly-L-lysine (PLL)[44]. For immobilisation using Ni^{2+}, 10 µL of 20mM HEPES, 3 mM $NiCl_2$, pH 7.4 solution was added to a freshly cleaved mica disk. Approximately 2 ng of DNA minicircles was added to the solution and adsorbed for 30 minutes. To remove any unbound DNA, the sample was washed four times using the same buffer solution. For immobilisation using PLL, 10 μL PLL (0.01% solution, MW 150,000-300,00; Sigma Aldrich) was deposited on the mica substrate and adsorbed for 1 min. The PLL surface was washed in a stream of MilliQ® ultrapure water, resistivity > 18.2 MΩ, and then washed four times with a 50 mM NaOAc pH 5.3 buffer solution to remove any PLL in solution. The supernatant was then removed and 10 μL 50 mM NaOAc pH 5.3 buffer solution was deposited on the surface. Approximately 2 ng of DNA minicircles was added to the solution and adsorbed for 30 minutes followed by four washes in the same buffer to remove any unbound DNA minicircles. Ni^{2+} immobilisation was used to obtain the data shown in Figures 1 and 2 and PLL for the data shown in Figure 3.

Triplex formation:

For experiments with TFO, DNA minicircles were incubated in an Eppendorf with a ten-fold excess of TFO in 50 mM NaOAc buffer at pH 5.3, prior to adsorption onto the mica substrate using the PLL method, as above. To verify the location of the TFO on the DNA sequence (Supp. Fig. 2), the minicircles were first linearised by cutting with NdeI.

AFM imaging:

All AFM measurements were performed in liquid. All experiments except Fig. 1f were carried out in PeakForce Tapping imaging on Multimode 8 and FastScan Bio AFM systems (Bruker). In these experiments, continuous force-distance curves were recorded with the tip-sample feedback set by the peak force as referenced to the force baseline. The following cantilevers were used: MSNL-E (Bruker) Peakforce HiResB (Bruker) and biolever mini (Olympus, Japan) on the Multimode 8, and FastScan D (Bruker) on the FastScan Bio with approximately equal resolution obtained by each. Force-distance curves were recorded over 20 nm (PeakForce Tapping amplitude of 10 nm), at frequencies of 4 (Multimode 8) and 8 (FastScan Bio) kHz. Imaging was carried out at PeakForce setpoints in the range of 5-20 mV, corresponding to peak forces of <70 pN. Images were recorded at 512 × 512 pixels to ensure a resolution ≥ 1 nm/pixel at line rates of 1-4 Hz.

Figure 1f was obtained on a home-built microscope with a closed-loop PicoCube XYZ piezo scanner (PhysikInstrumente, Karlsruhe, Germany) and with a Fabry-Perot interferometer to detect the cantilever deflection[19]. FastScan D (Bruker) cantilevers were actuated photothermally in tapping mode at amplitudes of 1-2 nm. Imaging was carried out at line rates of 3 Hz, over scan sizes of 50 nm with a setpoint ~80 % of the free amplitude, which correlates to a peak force of ≤100 pN[19].

AFM image processing:

AFM images were processed using a user-designed Python script (pygwytracing.py) which utilises the Gwyddion 'pygwy' module[37] for automated image correction, DNA molecule identification and morphological analysis. The algorithm searches recursively for files within a user-defined directory. This search also excludes any files of the format '_cs' which are cropped files exported by the Nanoscope Analysis software (Bruker, CA, USA). AFM images are loaded using gwyddion functions and topography data is automatically selected using the choosechannels function. The pixel size and dimensions of each image are determined using the imagedetails function, which allows all inputs to be specified in real, i.e. nanometre values, in place of pixel values. This is especially important for datasets with changing resolution.

Basic image processing is performed in the function editfile which uses the functions: 'align rows' to remove offsets between scan lines; 'level' to remove sample tilt as a first order polynomial; 'flatten base' which uses a combination of facet and polynomial levelling with automated masking; and 'zeromean' which sets the mean value of the image, i.e. the background, to zero. A gaussian filter (σ = 1.5) of 3.5 pixels (1-2 nm) was applied to remove pixel errors and high frequency noise.

Single DNA molecules are identified in images using a modified extension of Gwyddion's automated masking protocols, in which masks are used to define the positions of individual features (grains) on the imaged surface. The grains within a flattened AFM image are identified using the 'mask_outliers' function, which masks data points with height values that deviate from the mean by more than 1σ (with 3σ corresponding to a standard gaussian). Grains which touch the edge of the image (i.e. are incomplete) are removed using the 'grains_remove_touching_border' function and grains which are smaller than 200 nm^2 are removed using the 'grains_remove_by_size' function. Erroneous grains are removed using the *removelargeobjects* and *removesmallobjects* functions, which themselves use the function *find_median_pixel_area* to determine the size range of objects to remove. The 'grains_remove_by_size' function is then called again to remove grains which fall outside 50 % - 150 % of the median grain area determined in the previous step.

Grain statistics are then calculated for each image using the *grainanalysis* function which utilises the 'grains_get_values' function to obtain a number of statistical properties which are saved using the saveindividualstats function as '.json' and '.txt' files for later use in a subdirectory 'GrainStatistics' in the specified path. In addition, each grain's values are appended to an array [appended_data], to statistically analyse the morphologies of DNA molecules from all images for a given experiment (presumed to be within a single directory). This array is converted to a pandas dataframe[45] using the *getdataforallfiles* function and saved out using the savestats function as '.json' and '.txt' files with the name of the directory in the original path.

Individual grains (i.e. isolated molecules) are cropped out using the function *bbox*, which uses the grain centre x and y positions obtained in the *grainanalysis* function to duplicate the original image and crop it to a predefined size (here 80 nm) around the centre of the grain. These images are then labelled with the grain ID and saved out as tiff files in a subdirectory 'Cropped' in the specified path.
To allow for further processing in python, there is an option to obtain the image or mask as a numpy array[46], using the function *exportasnparray*. The processed image, and a copy with the mask overlaid are saved out using the *savefiles* function to a subdirectory 'Processed' in the specified path.
Statistical analysis and plotting is performed using the *statsplotting* script. This script uses the *importfromjson* function to import the JSON format file exported by *pygwytracing* and calculates various statistical parameters for all grain quantities, e.g. length, width and saves these out as a new JSON file using the *savestats* function. Both KDE plots and histograms are generated for any of the grain quantities using the matplotlib[47] and seaborn[48] libraries within the functions *plotkde*, *plotcolumns* and *plothist*.

Determination of minicircle bend angles by AFM:
To determine the bend angles for DNA minicircles by AFM, images were imported into Gwyddion, and basic processing was carried out as described above in the editfile script for basic flattening. Bend angles were then measured using Gwyddion's measurement tool.

Determination of triplex binding by AFM:
To verify that the small protrusions observed on DNA in the presence of the triplex forming oligonucleotide (TFO) at low pH were triplexes the site of the protrusions was determined. The 339 bp minicircles were linearised at the NdeI site and imaged by AFM as described above. Processed images were traced by hand in IMOD[2849] (University of Colorado, CO, USA) to determine the position of the protrusion along the DNA (Supp. Fig. 2). The tracing data was analysed using the *TFOlength* script. The mean and standard deviation for each length measurement (full minicircle, triplex, and triplex flanking lengths) were calculated using built in functions, and the data for each plotted as a histogram.

The length of the minicircle was determined as 109 ± 4 nm, with the triplex measured as 37 ± 2 nm, 34% of the length of the minicircle. The distance between the TFO site and the restriction site is 127 bp, which is 37 % of the length of the minicircle, and in good agreement with the AFM measurements. The length of the triplex as measured by AFM is 6 ± 2 nm. Errors quoted are standard deviations.

Atomistic simulations

Set up of the structures for supercoiled 339 bp DNA minicircles

Linear starting DNA molecules with the same 339 bp sequence as above were built using the NAB module implemented in AmberTools12[50]. DNA planar circles corresponding to six topoisomers (ΔLk = -6, -3, -2, -1, 0, 1) with/without the 16 bp triplex-forming oligomer were then constructed using an in-house programme. The AMBER99 forcefield[51] with different corrections for backbone dihedral angles including the parmBSC0 for α and γ[52], the parmOL4 for χ (glycosidic bond)[53] and the parmOL1 for ε and ζ[54] were used to describe the DNA. Parameters for protonated cytosine present in the triplex-forming oligomer were obtained from Soliva et al[55]. Following our standard protocol[56], the SANDER module within AMBER12 was used to subject the starting structures for the different type of minicircles to 20 ns of implicitly solvated MD using the Generalized Born/Solvent Accessible area (GB/SA) method[57] at 300 K and 200 mM salt concentration, with the long-range electrostatic cut-off set to 100. Restraints were imposed on the hydrogen bonding. Due to the neglect of solvent damping, the timescales in implicitly solvated MD are accelerated relative to simulations performed in solvent by at least 10 fold[13].

Simulations of 339bp minicircles in explicit solvent

To select the starting structure for explicitly solvated simulations, we performed clustering analysis using the average linkage algorithm within PTRAJ for the implicitly solvated DNA trajectories. Representative structures of the most populated clusters then were chosen and solvated in TIP3P rectangular boxes with a 6 nm buffer, 339 Ca^{2+} counterions[58] to balance the DNA charge and additional $Ca^{2+}/2Cl^{-}$ ion pairs[59] corresponding to a 100 mM. Two replicas of the -2 and -3 topoisomers were subjected to 100ns explicitly solvated MD simulations, starting from the two most representative structures. Single 100ns MD simulations were performed for topoisomers -6, -1, 0 and +1. Solvated MD runs were performed using the GROMACS 4.5 program[60] with standard MD protocols[56] at 308 K and, afterwards, were carefully visualised to ensure that rotation of the solute was not significant compared to the size of the simulation box over the timescale of the MD. Only the last 30 ns sampled every 10 ps were used for the subsequent analysis. VMD[61] and Chimera[62] were used to depict representative structures, to measure the longest distance across plectonemic loops and to detect defective DNA through visual inspection. DNA defects were confirmed through energetic analysis of stacking and hydrogen bonds at the relevant base steps using GROMACS 4.5. Hydrogen bonds were determined using 3.5 Å and 140^{0} as a distance and angle cutoff, respectively, as in Figure 4d.

Additional simulations for ΔLk =-6, -2 and 0 topoisomers were performed using BSC1 forcefield corrections[63] for DNA backbone dihedral angles instead of parmOL4. Simulations were started using same initial structures and were run with equivalent solvent conditions in TIP3P rectangular boxes with a 3 nm buffer for 100 ns using CUDA version of AMBER16[63]. Again, trajectories were carefully visualised to ensure that rotation of the solute was not significant compared to the size of the simulation box over the timescale of the MD[64]. Only the last 30 ns and a snapshot every 10 ps were used for the subsequent analysis, which are presented in Supp. Fig. 4

Simulations of linear DNA in explicit solvent

A 36-mer fragment containing the triplex-binding site (TBS) was extracted from the 339 bp minicircle to compare binding energies of this site on unconstrained linear DNA or on supercoiled minicircles. The TBS was placed in the middle to avoid end-effects[65]. The linear starting structure was solvated explicitly for running MD simulations and was set up, minimized, and equilibrated following the protocols described previously.

Simulations of 260 bp minicircles

The structure used for mirroring the high-resolution AFM image of a 251 bp minicircle on Figure 1a was extracted from a simulation previously run for the -1 topoisomer of a 260 bp minicircle[56]. The slightly longer sequence of 260 bp was constructed based on the experimental sequence of 251 bp studied here.

Global shape and other geometrical analysis of simulations

The radius of gyration was determined using the AMBER program PTRAJ (40). Other geometrical descriptions of the global shape, such as writhe and bend, were performed using *WrLINE* molecular contour[66] and *SerraLINE* program (both software suites are freely accessible at https://agnesnoylab.wordpress.com/software/). With *SerraLINE*, the bending angles θ were calculated from the directional correlation, $\theta = \cos^{-1}(z_i \cdot z_j)$, where z_i and z_j are the two tangent vectors. Each z_i was obtained by combining two successive points of the *WrLINE* global contour ($z_i = r_{i+1} - r_i$). Bending angles θ were calculated using two tangent vectors (z_i and z_j) separated by 16 nucleotides (approximately a DNA helical turn and a half) as a compromise length for capturing the overall bend produced by a defect or by canonical B-DNA. The bending profiles in Supp. Fig. 1 were obtained by scanning all the possible 16-bp sub-fragments along the minicircle and the peaks were selected to compare the MD simulations with the AFM data shown in Figure 2d.

Energy calculations of triplex DNA formation

To obtain theoretical insight into the thermodynamics driving triplex binding, we used the MD trajectories to estimate the global (e.g. electrostatics) and local (e.g. base-pair stacking and hydrogen bonding) contributions to the overall binding energy. The global electrostatic contribution for configurational energy (E_{ele}) was evaluated using the AMBER program MMPBSA[67]. To compare between the different topoisomers, the individual components were referred to the relaxed DNA-naked topoisomer ($\Delta E_{elec;\,0} = E_{elec} - E_{elec;0}$) as it is shown in Figure 4c.

We also analysed the interaction energy between nucleotides in the triplex binding site, considering in-plane base interactions and nearest neighbours only (e.g. 9 bases in total). All interaction energies were calculated using the GROMACS 4.5 program. The two components of the binding energy were calculated at the TBS for each topoisomer by discarding the effect of the unbound 3rd strand ($E_{bind} = E_{TRI} - E_{DNA}$). Values for the different topoisomers were referenced to the linear fragment ($\Delta E_{bind;\,L} = E_{bind} - E_{bind;L}$), as shown in Figure 4d. The local interaction energy terms (Lennard-Jones and electrostatic interactions) between in plane nucleotides were used as an estimate of the Hoogsteen hydrogen bonding at the triplex binding site (TBS) (Fig. 4d, yellow bars), while the interaction energies between bases in the planes above and below were used as an estimation of base-stacking, hydrogen bonding and non-bonded backbone interactions (4d, blue bars). The presence of these hydrogen bonds was confirmed by visual inspection in VMD (Fig. 4b).

Surface plasmon resonance (SPR)

SPR measurements were recorded at either 25°C or 35°C using a Biacore T200 system (GE Healthcare). All experiments were performed using an SA Series S Sensor Chip (GE Healthcare), which has four flow cells each containing streptavidin pre-immobilised to a carboxymethylated dextran matrix. For immobilisation, a standard immobilisation protocol was used with a running buffer of HSB-EP+ buffer (10 mM HEPES pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.05% [v/v] surfactant P20). The chip surface was first washed using three injections of 1.0 M NaCl, 50 mM NaOH for 60 s, each followed by buffer for 60 s (all at 10 µL/min). The 5´-biotinylated triplex-forming oligonucleotide (TFO1R, 30-60 nM) was then immobilised onto two of the flow cells (FC2 and FC4) and a response of approximately 250 Response Units (RU) was aimed for. The remaining two flow cells (FC1 and FC3) were kept free of ligand and were used as reference cells.

Experiments were carried out using conditions that were modified from those reported previously with plasmid pNO1[16], optimised for the 339 minicircle. Using these optimised conditions (TFO1R [250 RU immobilised], flow rate 2 μL min^{-1}, 100 mM calcium acetate pH 4.8, 25°C, injection time 600 s; regeneration: 1 M NaCl, 5 mM NaOH, 60 s, 30μL min^{-1}) a range of differently supercoiled samples at 50

nM of 339 nr (ΔLk -4.9 to +1) as well as relaxed, nicked, linear, and samples containing no triplex-forming sequences were injected and the binding monitored.

The kinetics of the binding between the small-circle DNA substrate and the TFO1R ligand were then measured using a multi-cycle kinetics approach using the same optimised conditions but with 221 RU of TFO1R immobilised. For the kinetic experiments, 339 bp minicircles (ΔLk -4.9, -2.8, linear and relaxed) were injected over flow cells 1 and 2 for 600 s at a range of concentrations (2.5, 5, 10, 20, 30 and 40 nM) and a buffer-only control. A buffer only solution was then flowed for 1 h so that the dissociation could be more accurately recorded. The SA chip was regenerated after each injection of DNA using 1 M NaCl, 5 mM NaOH. The experiment was carried out at 35°C with a flow rate of 2 μLmin^{-1} using 100 mM calcium acetate pH 4.8 as the running buffer. The inclusion of buffer-only controls enabled the use of double referencing, whereby, for each analyte measurement, in addition to subtracting the response in the reference flow cells from the response in the test flow cells, a further buffer-only subtraction was made to correct for the bulk refractive index changes or machine effects[68]. The data were analysed using the Biacore T200 Evaluation software version 2.0 using the kinetics fit assuming a 1:1 binding model.

References:

1. Bates, A. D. & Maxwell, A. *DNA Topology.* (Oxford University Press, 2005).
2. Fogg, J. M. *et al.* Bullied no more: When and how DNA shoves proteins around. *Q. Rev. Biophys.* **45**, 257–299 (2012).
3. Mirkin, S. M. *DNA Topology: Fundamentals*. (John Wiley & Sons, Ltd, 2001).
4. Schvartzman, J. B., Hernández, P., Krimer, D. B., Dorier, J. & Stasiak, A. Closing the DNA replication cycle: from simple circular molecules to supercoiled and knotted DNA catenanes. *Nucleic Acids Res.* (2019).
5. Higgins, N. P. & Vologodskii, A. V. Topological Behavior of Plasmid DNA. *Microbiol. Spectr.* **3**, (2015).
6. Sobetzko, P. Transcription-coupled DNA supercoiling dictates the chromosomal arrangement of bacterial genes. *Nucleic Acids Res.* **44**, 1514–1524 (2016).
7. Levens, D. Cellular MYCro Economics: Balancing MYC Function with MYC Expression. *Cold Spring Harb. Perspect. Med.* **3**, a014233–a014233 (2013).
8. Travers, A. & Muskhelishvili, G. A common topology for bacterial and eukaryotic transcription initiation? *EMBO Rep.* **8**, 147–151 (2007).
9. van Aelst, K., Martínez-Santiago, C. J., Cross, S. J. & Szczelkun, M. D. The Effect of DNA Topology on Observed Rates of R-Loop Formation and DNA Strand Cleavage by CRISPR Cas12a. *Genes* **10**, 169 (2019).
10. Newton, M. D. *et al.* DNA stretching induces Cas9 off-target activity. *Nat. Struct. Biol.* **32**, 1 (2019).

11. Benham, C. J., Brady, G. W. & Fein, D. B. X-ray scattering from randomly oriented superhelices. Circular superhelical DNA. *Biophys. J.* **29**, 351–366 (1980).
12. Bendel, P. & James, T. L. Structural and dynamic differences between supercoiled and linear DNA from proton NMR. *Proc. Natl. Acad. Sci.* **80**, 3284–3286 (1983).
13. Irobalieva, R. N. *et al.* Structural diversity of supercoiled DNA. *Nat. Commun.* **6**, 8440 (2015).
14. Koster, D. A., Crut, A., Shuman, S., Bjornsti, M.-A. & Dekker, N. H. Cellular Strategies for Regulating DNA Supercoiling: A Single-Molecule Perspective. *Cell* **142**, 519–530 (2010).
15. Hanvey, J. C., Shimizu, M. & Wells, R. D. Intramolecular DNA triplexes in supercoiled plasmids. *Proc. Natl. Acad. Sci. U. S. A.* **85**, 6292–6296 (1988).
16. Maxwell, A., Burton, N. P. & O'Hagan, N. High-throughput assays for DNA gyrase and other topoisomerases. *Nucleic Acids Res.* **34**, e104–e104 (2006).
17. Frank-Kamenetskii, M. D. & Mirkin, S. M. Triplex DNA structures. *Annu. Rev. Biochem.* **64**, 65–95 (1995).
18. Curuksu, J., Zacharias, M., Lavery, R. & Zakrzewska, K. Local and global effects of strong DNA bending induced during molecular dynamics simulations. *Nucleic Acids Res.* **37**, 3766–3773 (2009).
19. Leung, C. *et al.* Atomic Force Microscopy with Nanoscale Cantilevers Resolves Different Structural Conformations of the DNA Double Helix. *Nano Lett.* **12**, 3846–3850 (2012).
20. Ido, S. *et al.* Beyond the Helix Pitch: Direct Visualization of Native DNA in Aqueous Solution. *ACS Nano* **7**, 1817–1822 (2013).
21. Pyne, A., Thompson, R., Leung, C., Roy, D. & Hoogenboom, B. W. Single-Molecule Reconstruction of Oligonucleotide Secondary Structure by Atomic Force Microscopy. *Small* **10**, 3257–3261 (2014).
22. Kim, S. H. *et al.* DNA sequence encodes the position of DNA supercoils. *eLife* **7**, e36557 (2018).
23. Shoura, M. J. *et al.* Intricate and Cell-type-specific Populations of Endogenous Circular DNA (eccDNA) in Caenorhabditis elegans and Homo sapiens. *G3 GenesGenomesGenetics* **7**, 3295 (2017).
24. Møller, H. D. *et al.* Circular DNA elements of chromosomal origin are common in healthy human somatic tissue. *Nat. Commun.* **9**, 1069 (2018).
25. Paulsen, T., Shibata, Y., Kumar, P., Dillon, L. & Dutta, A. Small extrachromosomal circular DNAs, microDNA, produce short regulatory RNAs that suppress gene expression independent of canonical promoters. *Nucleic Acids Res.* **47**, 4586–4596 (2019).
26. Brouns, T. *et al.* Free Energy Landscape and Dynamics of Supercoiled DNA by High-Speed Atomic Force Microscopy. *ACS Nano* **12**, 11907–11916 (2018).
27. Lee, A. J., Szymonik, M., Hobbs, J. K. & Wälti, C. Tuning the translational freedom of DNA for high speed AFM. *Nano Res.* **8**, 1811–1821 (2015).

28. Langowski, J., Kapp, U., Klenin, K. & Vologodskii, A. Solution structure and dynamics of DNA topoisomers: Dynamic light scattering studies and Monte Carlo simulations. *Biopolymers* **34**, 639–646 (1994).

29. Matek, C., Ouldridge, T. E., Doye, J. P. K. & Louis, A. A. Plectoneme tip bubbles: Coupled denaturation and writhing in supercoiled DNA. *Sci. Rep.* **5**, 7655 (2015).

30. Zhabinskaya, D. & Benham, C. J. Theoretical Analysis of Competing Conformational Transitions in Superhelical DNA. *PLoS Comput. Biol.* **8**, e1002484 (2012).

31. Scott, S. *et al.* Visualizing structure-mediated interactions in supercoiled DNA molecules. *Nucleic Acids Res.* **14**, 2206 (2018).

32. Oberstrass, F. C., Fernandes, L. E. & Bryant, Z. Torque measurements reveal sequence-specific cooperative transitions in supercoiled DNA. *Proc. Natl. Acad. Sci.* **109**, 6106–6111 (2012).

33. Lyubchenko, Y. L. & Shlyakhtenko, L. S. Visualization of Supercoiled DNA with Atomic Force Microscopy in situ. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 496–501 (1997).

34. Liu, Y., Bondarenko, V., Ninfa, A. & Studitsky, V. M. DNA supercoiling allows enhancer action over a large distance. *Proc. Natl. Acad. Sci.* **98**, 14883–14888 (2001).

35. Janićijević, A. *et al.* DNA bending by the human damage recognition complex XPC–HR23B. *DNA Repair* **2**, 325–336 (2003).

36. Desai, P. R., Das, S. & Neuman, K. C. Coarse-Grained Modeling of DNA Plectoneme Formation in the Presence of Base-Pair Mismatches. *Biophys. J.* **118**, 221a (2020).

37. Nečas, D. & Klapetek, P. Gwyddion: an open-source software for SPM data analysis. *Cent. Eur. J. Phys.* **10**, 181–188 (2011).

38. Chan, P. P. & Glazer, P. M. Triplex DNA: fundamentals, advances, and potential applications for gene therapy. *J. Mol. Med.* **75**, 267–282 (1997).

39. Hatfield, G. W. & Benham, C. J. DNA topology-mediated control of global gene expression in Escherichia coli. *Annu. Rev. Genet.* **36**, 175–203 (2002).

40. McGuffie, E. M., Pacheco, D., Carbone, G. M. & Catapano, C. V. Antigene and antiproliferative effects of a c-myc-targeting phosphorothioate triple helix-forming oligonucleotide in human leukemia cells. *Cancer Res.* **60**, 3790–3799 (2000).

41. Ronald, J. A., Chuang, H.-Y., Dragulescu-Andrasi, A., Hori, S. S. & Gambhir, S. S. Detecting cancers through tumor-activatable minicircles that lead to a detectable blood biomarker. *Proc. Natl. Acad. Sci.* **112**, 3068–3073 (2015).

42. Zhao, N., Fogg, J. M., Zechiedrich, L. & Zu, Y. Transfection of shRNA-encoding Minivector DNA of a few hundred base pairs to regulate gene expression in lymphoma cells. *Gene Ther.* **18**, 220–224 (2011).

43. Fogg, J. M. *et al.* Exploring writhe in supercoiled minicircle DNA. *J. Phys. Condens. Matter* **18**, S145–S159 (2006).

44. Pyne, A. L. B. & Hoogenboom, B. W. Imaging DNA Structure by Atomic Force Microscopy. in *Chromosome Architecture* (ed. Leake, M. C.) 47–60 (Springer New York, 2016).

45. McKinney, W. Data structures for statistical computing in python. in *Proceedings of the 9th Python in Science Conference* vol. 445 51–56 (Austin, TX, 2010).

46. Van Der Walt, S., Colbert, S. C. & Varoquaux, G. The NumPy array: a structure for efficient numerical computation. *Comput. Sci. Eng.* **13**, 22 (2011).

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

48. Waskom, M. *et al. Seaborn: V0.5.0 (November 2014).* (Zenodo, 2014). doi:10.5281/ZENODO.12710.

49. Kremer, J. R., Mastronarde, D. N. & McIntosh, J. R. Computer Visualization of Three-Dimensional Image Data Using IMOD. *J. Struct. Biol.* **116**, 71–76 (1996).

50. Cheatham, T. *et al.* AMBER 11. *Univ. Calif.* (2010).

51. Cheatham, T. E., Cieplak, P. & Kollman, P. A. A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. *J. Biomol. Struct. Dyn.* **16**, 845–862 (1999).

52. Pérez, A. *et al.* Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers. *Biophys. J.* **92**, 3817–3829 (2007).

53. Krepl, M. *et al.* Reference simulations of noncanonical nucleic acids with different χ variants of the AMBER force field: quadruplex DNA, quadruplex RNA and Z-DNA. *J. Chem. Theory Comput.* **8**, 2506–2520 (2012).

54. Zgarbová, M. *et al.* Toward Improved Description of DNA Backbone: Revisiting Epsilon and Zeta Torsion Force Field Parameters. *J. Chem. Theory Comput.* **9**, 2339–2354 (2013).

55. Soliva, R., Laughton, C. A., Luque, F. J. & Orozco, M. Molecular Dynamics Simulations in Aqueous Solution of Triple Helices Containing d(G·C·C) Trios. *J. Am. Chem. Soc.* **120**, 11226–11233 (1998).

56. Noy, A., Maxwell, A. & Harris, S. A. Interference between Triplex and Protein Binding to Distal Sites on Supercoiled DNA. *Biophys. J.* **112**, 523–531 (2017).

57. Tsui, V. & Case, D. A. Theory and applications of the generalized Born solvation model in macromolecular simulations. *Biopolymers* **56**, 275–291 (2000).

58. Bradbrook, G. M. *et al.* X-Ray and molecular dynamics studies of concanavalin-A glucoside and mannoside complexes Relating structure to thermodynamics of binding. *J. Chem. Soc. Faraday Trans.* **94**, 1603–1611 (1998).

59. Smith, D. E. & Dang, L. X. Computer simulations of NaCl association in polarizable water. *J. Chem. Phys.* **100**, 3757–3766 (1994).

60. Hess, B., Kutzner, C., van der Spoel, D. & Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **4**, 435–447 (2008).

61. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
62. Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
63. Ivani, I. *et al.* Parmbsc1: a refined force field for DNA simulations. *Nat. Methods* **13**, 55–58 (2016).
64. Case, D. *et al.* Kollman Amber 16. *San Franc. Univ. Calif.* (2016).
65. Noy, A. & Golestanian, R. Length Scale Dependence of DNA Mechanical Properties. *Phys. Rev. Lett.* **109**, 228101 (2012).
66. Sutthibutpong, T., Harris, S. A. & Noy, A. Comparison of Molecular Contours for Measuring Writhe in Atomistic Supercoiled DNA. *J. Chem. Theory Comput.* **11**, 2768–2775 (2015).
67. Miller, B. R. *et al.* MMPBSA.py: An Efficient Program for End-State Free Energy Calculations. *J. Chem. Theory Comput.* **8**, 3314–3321 (2012).
68. Myszka, D. G. Improving biosensor analysis. *J. Mol. Recognit.* **12**, 279–284 (1999).

Acknowledgements

We thank Lynn Zechiedrich and Jonathan Fogg for informative conversations and for supplying minicircle plasmids; Andrea Slade, James Shaw, Bede Pittenger, Shuiqing Hu, Chanmin Su and Thomas Mueller (Bruker) for assistance in developing equipment and protocols for high-resolution imaging of DNA; Maya Topf, Agnel Joseph, Joanna Lis and Christopher Soelistyo for assistance with developing the automated AFM analysis; Andrea Hall for assistance with fermentation; and Charlie Laughton for the program to build *in silico* DNA minicircles with triplex DNA. This work was supported by grants BB/I019294/1 and BB/J004561/1 from the Biotechnology and Biological Sciences Research Council (BBSRC), by grants EP/M506448/1, EP/M028100/1 and EP/N027639/1, EP/R513143/1 from the Engineering and Physical Sciences Research Council (EPSRC), by a UKRI/MRC Rutherford Innovation fellowship MR/R024871, and by the John Innes Foundation. Time on ARCHER and JADE was granted via the UK High-End Computing Consortium for Biomolecular Simulation, HECBioSim, supported by the EPSRC (grant No. EP/R029407/1).

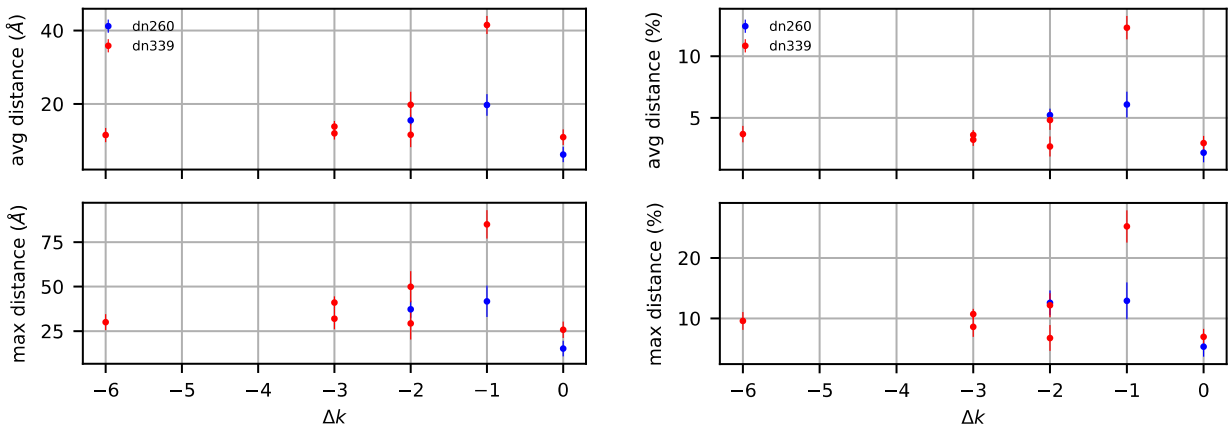
Author Contributions

A.L.B.P., A.N., B.W.H., A.D.B, A.M., and S.A.H. conceived and designed the experiments. A.L.B.P. and K.M. conducted AFM experiments, A.L.B.P. and J.G.B. wrote software to analyse AFM images and performed analysis. A.N. performed and analysed MD simulations and V.V. developed SerraLINE program. F.M.C., M.M.P. and C.E.M. performed triplex-binding SPR experiments. A.L.B.P., L.A.M., F.M.C., M.M.P performed gel electrophoresis and prepared the different samples. A.L.B.P, A.N. and S.A.H analysed the data and wrote the manuscript with input from B.W.H, A.D.B., A.M.

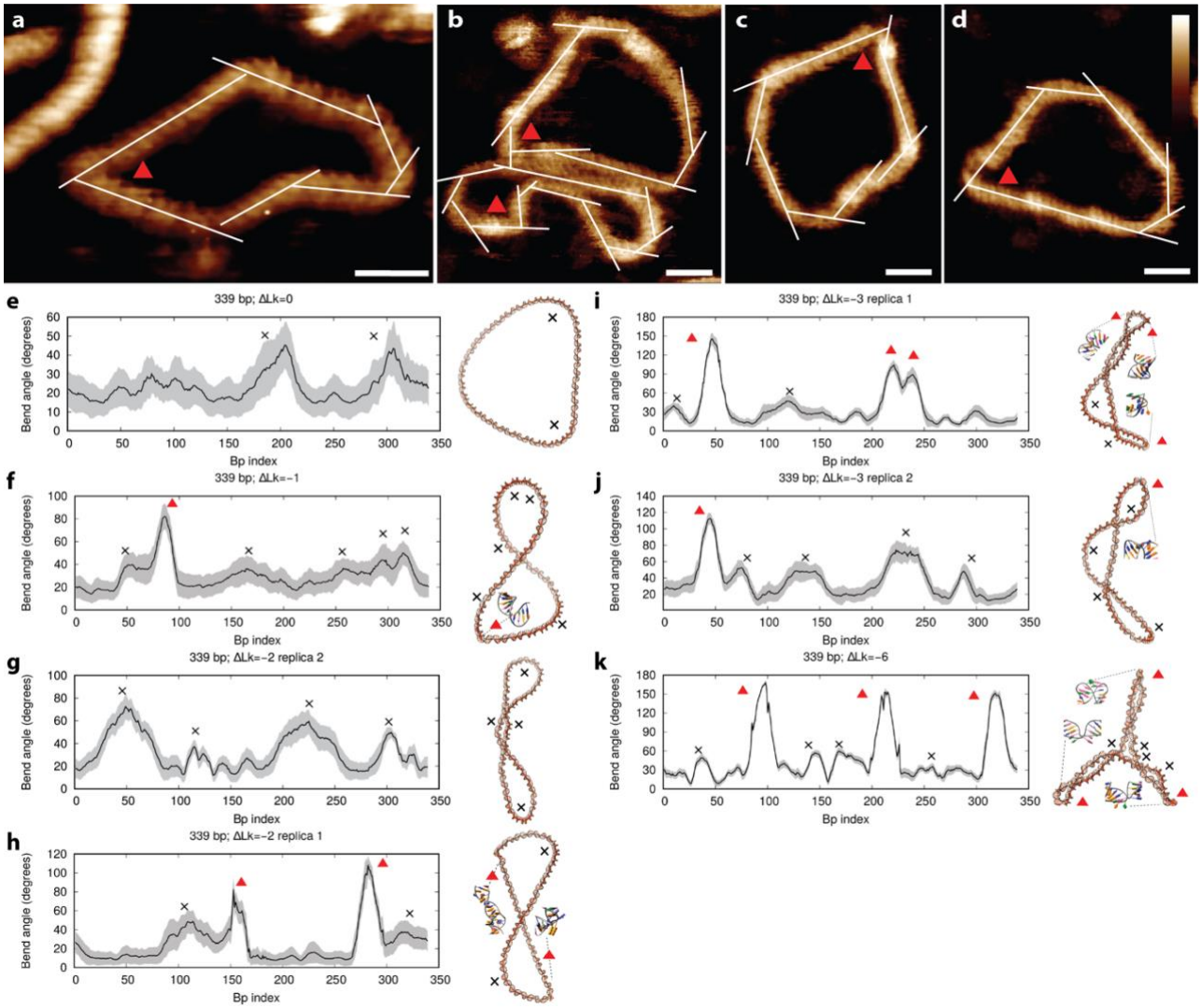
Base-pair resolution analysis of the effect of supercoiling on DNA flexibility and recognition

Alice L. B. Pyne, Agnes Noy, Kavit Main, Victor Velasco-Berrelleza, Michael M. Piperakis, Lesley A. Mitchenall, Fiorella M. Cugliandolo, Joseph G. Beton, Clare E.M. Stevenson, Bart W. Hoogenboom, Andrew D. Bates, Anthony Maxwell, Sarah A. Harris

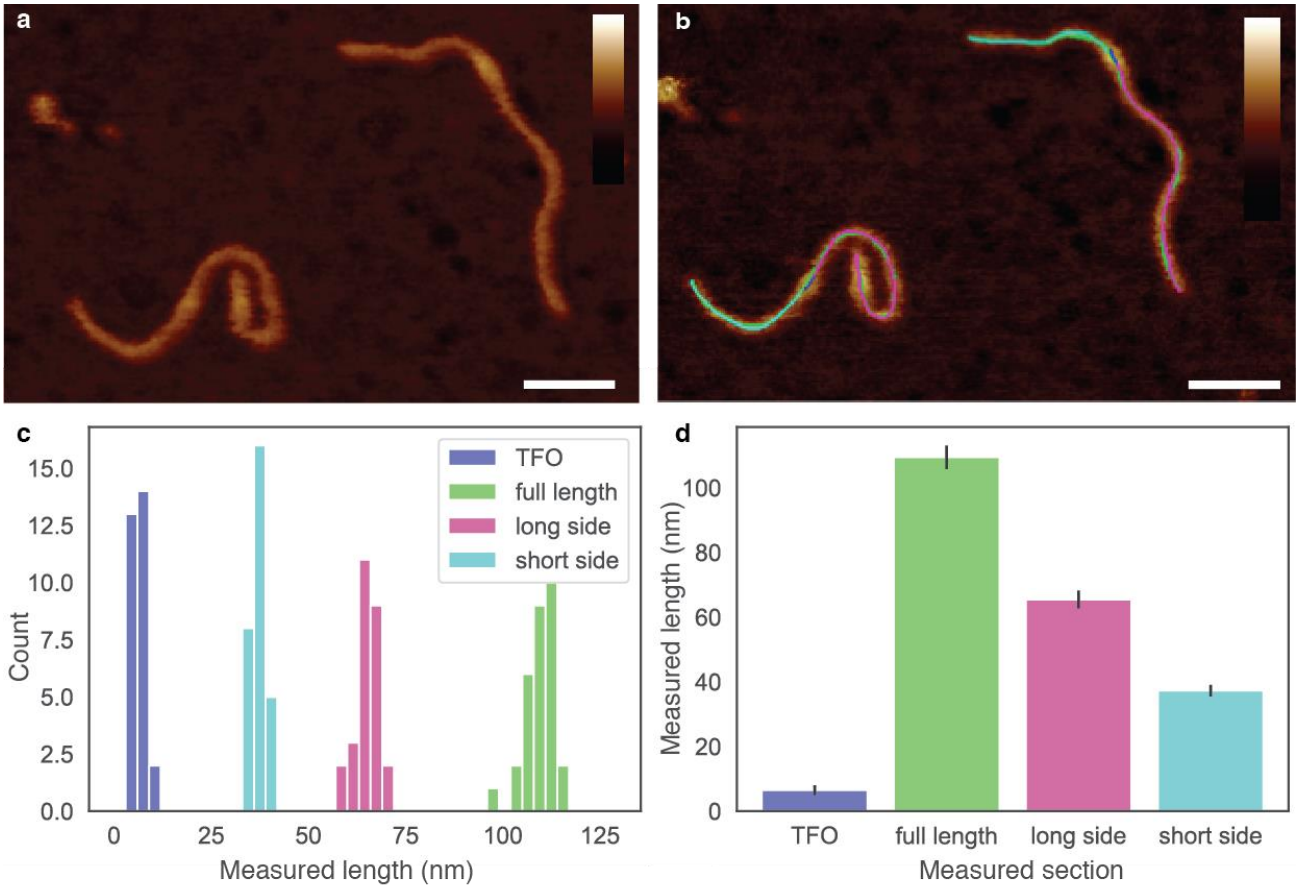
Supplementary Information



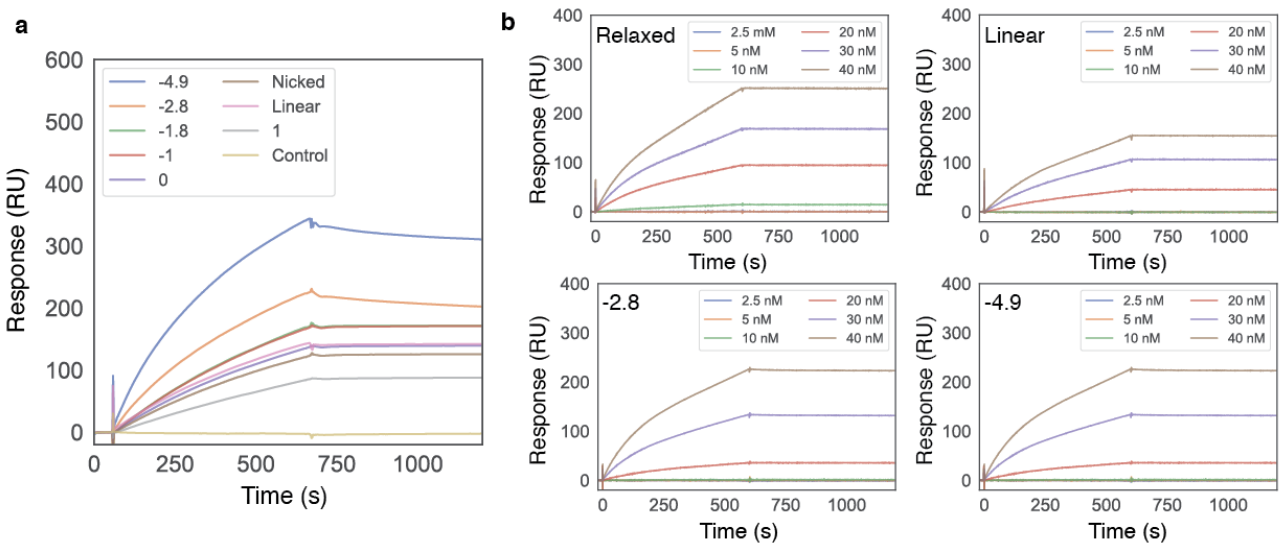
Supplementary Figure 1: In silico analysis shows a high degree of planarity in both 260 bp (dn260) and 339 bp (dn339) minicircles. For both systems, the average deviation from planarity is ≤ 2 nm (~4%) for all superhelical densities except ΔLk = -1 where for the 339 bp molecule this rises to 4 nm (~12%). The maximum deviation from planarity at any one point is less than 25%.



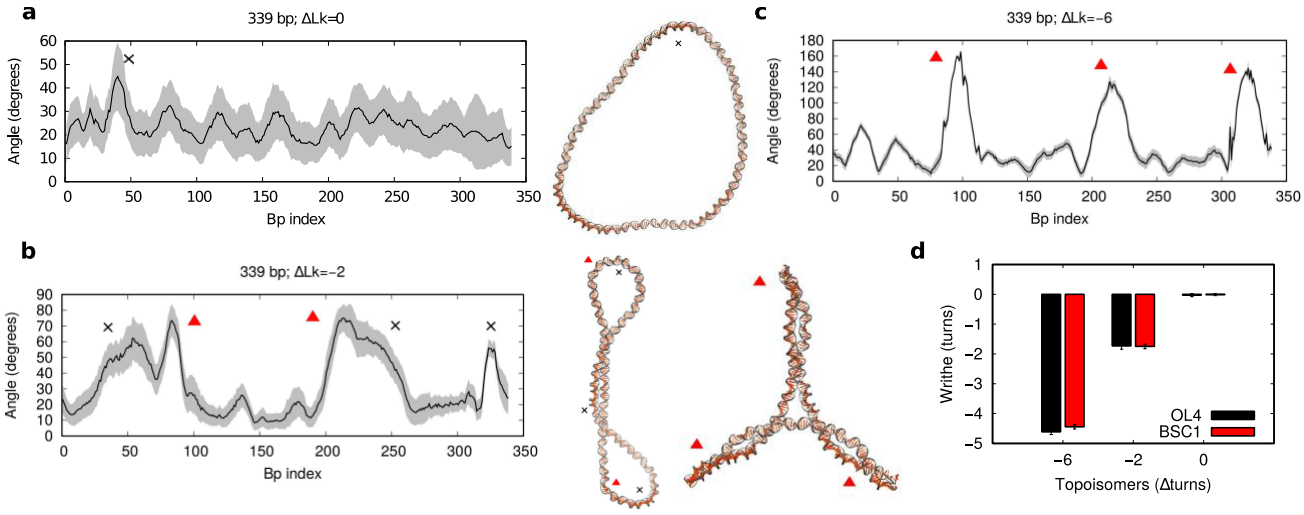
Supplementary Figure 2: Analysis of bend angles by AFM and MD simulations. **a**-**d** High resolution AFM images of 251 (**a**) and 339 (**b**, **c**, **d**) bp minicircles, analysed using Gwyddion[1] to calculate bend angles. The bend angle for bent and kinked regions were calculated as 51 ± 14° and 106 ± 16°. Errors quoted are standard deviation (N = 33 and 5 respectively). At this resolution the minicircles show observable differences in structure, as kinked (red triangle) or bent regions. Scale bars 10 nm, height scale (inset) 2.5 nm. **e-k** Bend angles for all sub-fragments 16 bp-long of MD simulations calculated by the SerraLINE program using the WrLINE profile. All peaks higher than 40º are selected for Figure 2d and are classified as B-DNA bends (black cross) or defects (red triangles) depending on whether canonical non-bonded interactions were broken or not. Grey shades correspond to standard deviations calculated along the last 30 ns of the simulations



Supplementary Figure 3: Confirmation of triplex formation in 339 bp minicircles by AFM. **a**, AFM image of linearised 339-bp minicircles showing a protrusion at a location consistent with triplex formation. Minicircles linearised with NdeI, incubated with the triplex forming oligonucleotide TFO1 and immobilised on mica using Poly-L-lysine in 50 mM NaOAc, pH 5.2, at room temperature. **b**, AFM images are hand traced in IMOD$_2$ to determine the length of the linearised minicircle, and the lengths of DNA either side of the protrusion. **c**, **d**, The length of the TFO, full length, long and short sides were determined by AFM to be 6 ± 2, 109 ± 4, 65 ± 3 and 37 ± 2 nm (SD) respectively. This is in good agreement with calculations by base pair separation for the full length and short sides as 115 nm (339 bp) 43 nm (127 bp). Scale bars: 25 nm, height scale (scale bar inset): 4 nm.



Supplementary Figure 4: Supercoiling does not substantially affect the kinetics of triplex formation in minicircle DNA. **a**, SPR shows capture of 339 bp minicircles (ΔLk -4.9 to +1, and relaxed (R), nicked (N), linear (L) and non-triplex forming 339 bp mincircle circle (C) by TFO1R. **b**, Kinetics of the triplex binding interaction. To determine the kinetics of this interaction a range of TFO1 concentrations (2.5, 5, 10, 20, 30 and 40 nM) were injected for each sample for 10 mins and then buffer flowed for 1 hour (for clarity sensorgrams are only shown for 20 mins). Kinetic parameters were extracted from **b**, shown in Supplementary Table 1. Biotinylated TFO1R was immobilised to 250 RU on a streptavidin chip in a Biacore T200 instrument and minicircles (50 nM) were flowed over the triplex forming sequence TFO1, immobilised on a streptavidin chip surface (2 µL/min, 600 s) in 100 mM calcium acetate, pH 4.8, 35°C.

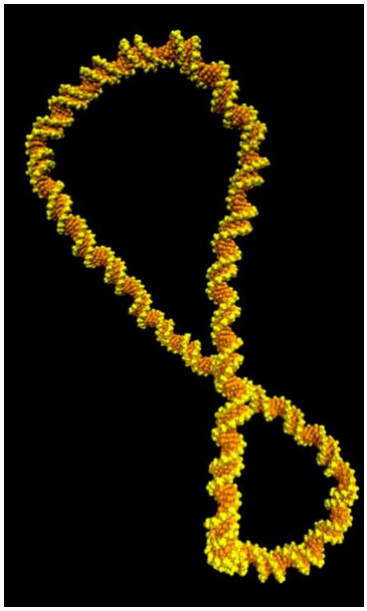


Supplementary Figure 5: OL4 and BSC1 DNA forcefields describe supercoiled DNA minicircles broadly with same structure and dynamics. **a-c** Average structures and bend angles obtained from BSC1 simulations are similar to the ones obtained from OL4 simulations (see Supplementary Figure 1). Bending profiles are calculated by SerraLINE program using the WrLINE profile for all sub-fragments 16 bp-long. All peaks higher than 40º are classified as B-DNA bends (black cross) or defects (red triangles) depending

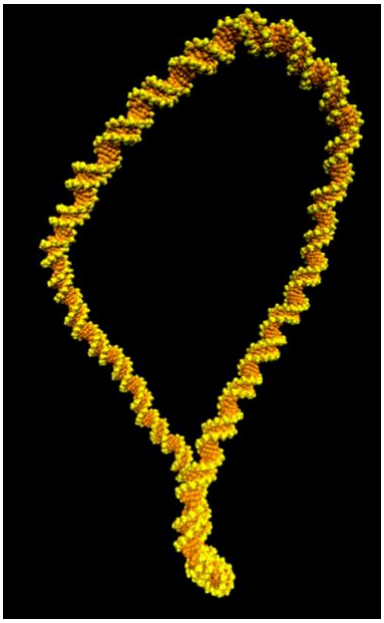
on whether canonical non-bonded interactions were broken or not. **d** Comparison of Writhe calculations obtained from the different forcefields. Grey shades and error bars correspond to standard deviations calculated along the last 30 ns of the simulations

339 nr minicircle ΔLK	k_{on} (ms^{-1}) [SE]	k_{off} (s^{-1}) [SE]	K_d (M)
-4.9	8.1×10^{4} [71]	3.4×10^{-6} [2.1×10^{-8}]	4.2×10^{-11}
-2.8	8.3×10^{4} [42]	1.9×10^{-6} [9.3×10^{-9}]	2.3×10^{-11}
Relaxed	6.2×10^{4} [91]	4.5×10^{-6} [2.7×10^{-8}]	7.3×10^{-11}
Linear	4.8×10^{4} [14]	3.6×10^{-7} [1.8×10^{-10}]	7.6×10^{-12}
No TFS -1.8	NB	NB	NB

Supplementary Table 1: Kinetics of triplex formation between TFO1R and 339 bp minicircle samples as measured by SPR. DNA analyte was flowed at 2 µL/min over a streptavidin-coated chip that had previously been loaded with TFO1R (221 RU). Running buffer was 100 mM calcium acetate pH 4.8 and all experiments were carried out at 35°C. The minicircle was injected at a range of concentrations from 2.5-40 nM for 600 s and the dissociation was measured for 1 h. Even with these parameters the dissociation was too slow to be accurately determined by the instrument. All data were fitted using the Biacore T200 evaluation software assuming a 1:1 binding model. The values obtained are broadly similar showing that the degree of supercoiling does not play a large effect in triplex formation when minicircles are used. (TFS = triplex-forming sequence; NB = no binding.)



Supplementary Video 1: A trajectory fragment 400 picosecond (ps) long extracted from the implicitly-solvated MD simulation of the 339 bp ΔLk -1 topoisomer mirroring the dynamics observed by AFM in Figure 1f (top sequence of structures). A series of four snapshots extracted from this trajectory equally-spaced at 100 ps are sketched on the top row of Figure 1g. Note that timescales are not physically meaningful because water viscosity is disregarded in the used model of implicit representation of the solvent.



Supplementary Video 2: A second trajectory fragment 300 ps long extracted from the implicitly-solvated MD simulation of the 339 bp ΔLk -1 topoisomer mirroring the dynamics observed by AFM in Figure 1f (bottom sequence of structures). A series of four snapshots extracted from this trajectory equally-spaced 75 ps are sketched on the bottom row of Figure 1g. As before, timescales are not physically meaningful.

References:

1. Nečas, D. & Klapetek, P. Gwyddion: an open-source software for SPM data analysis. *Cent. Eur. J. Phys.* **10**, 181–188 (2011).
2. Kremer, J. R., Mastronarde, D. N. & McIntosh, J. R. Computer Visualization of Three-Dimensional Image Data Using IMOD. *J. Struct. Biol.* **116**, 71–76 (1996).