

Coarse-grained modeling reveals the impact of supercoiling and loop length in DNA looping kinetics

Charles H. Starr,¹ Zev Bryant,^{1,2} and Andrew J. Spakowitz^{1,3,4,5,*}

¹Biophysics Program, Stanford University, Stanford, California; ²Department of Bioengineering, Stanford University, Stanford, California;

³Department of Chemical Engineering, Stanford University, Stanford, California; ⁴Department of Materials Science and Engineering, Stanford University, Stanford, California; and ⁵Department of Applied Physics, Stanford University, Stanford, California

ABSTRACT Measurements of protein-mediated DNA looping reveal that in vivo conditions favor the formation of loops shorter than those that occur in vitro, yet the precise physical mechanisms underlying this shift remain unclear. To understand the extent to which in vivo supercoiling may explain these shifts, we develop a theoretical model based on coarse-grained molecular simulation and analytical transition state theory, enabling us to map out looping energetics and kinetics as a function of two key biophysical parameters: superhelical density and loop length. We show that loops on the scale of a persistence length respond to supercoiling over a much wider range of superhelical densities and to a larger extent than longer loops. This effect arises from a tendency for loops to be centered on the plectonemic end region, which bends progressively more tightly with superhelical density. This trend reveals a mechanism by which supercoiling favors shorter loop lengths. In addition, our model predicts a complex kinetic response to supercoiling for a given loop length, governed by a competition between an enhanced rate of looping due to torsional buckling and a reduction in looping rate due to chain straightening as the plectoneme tightens at higher superhelical densities. Together, these effects lead to a flattening of the kinetic response to supercoiling within the physiological range for all but the shortest loops. Using experimental estimates for in vivo superhelical densities, we discuss our model's ability to explain available looping data, highlighting both the importance of supercoiling as a regulatory force in genetics and the additional complexities of looping phenomena in vivo.

SIGNIFICANCE How living cells are able to form DNA loops significantly shorter than those that occur in vitro represents a long-standing conundrum in molecular biophysics. Our theoretical model quantifies the extent to which DNA supercoiling is responsible for this discrepancy while also revealing a structural explanation for this effect. This work provides key physical insights into the loop formation process and sets the stage for future modeling aimed at more fully describing genome dynamics in vivo. Finally, our quantitative predictions guide the use of both superhelical density and loop length as control parameters in synthetic genetic systems.

INTRODUCTION

The formation of a DNA loop, in which a protein physically bridges two DNA sites along the same double helix, is a ubiquitous event in genetics with central roles in transcriptional control, recombination, replication, and chromosome organization (1,2). The dependence of looping efficiency on loop length is an extensively studied topic, and measurements of this dependence have led to key insights into the

physical factors that govern loop formation (3). For example, short, stiff DNA loops display oscillations in looping efficiency as a function of loop length, with the period equal to the helical repeat of DNA (~ 10.5 bp)—a reflection of the energetic cost associated with twisting DNA into helical alignment with its protein binding sites.

Across longer length scales, the mechanics of DNA bending control the looping process. Bare DNA in vitro is well described by the wormlike chain (WLC) model, which treats DNA as a continuous thread subject to a quadratic bending penalty and thermal fluctuations. Using the widely accepted bending persistence length $l_p \approx 50$ nm ≈ 150 bp for double-stranded DNA (dsDNA), the WLC model

Submitted September 13, 2021, and accepted for publication April 6, 2022.

*Correspondence: ajspakow@stanford.edu

Editor: Wilma K. Olson.

<https://doi.org/10.1016/j.bpj.2022.04.009>

© 2022 Biophysical Society.



predicts that loops in the ~ 400 – 500 bp range form most readily, with shorter and longer loops being relatively less probable due to enthalpic and entropic costs, respectively (4,5). This prediction is supported by a number of in vitro looping assays (6–9).

On the other hand, in vivo measurements show that the peak loop length is shifted toward shorter values (7,9,10), demonstrating that in vivo looping is governed by physical factors beyond those captured by the WLC model with $l_p = 50$ nm. In contrast to purified DNA in vitro, genomic DNA in living cells contains topological domains and bound architectural proteins, and can undergo active processes driven by molecular motors, all of which can have significant consequences on the looping reaction. However, the precise factors responsible for the in vitro-in vivo looping shifts are not fully understood.

In the face of such complexity, a quantitative dissection of each factor in isolation is one route toward unifying the observed looping data, with the added advantage that physical principles we reveal provide a basis for new levels of control over synthetic genetic circuits, many of which already utilize DNA looping systems (11–15). In this article we specifically model the effect of supercoiling on loop formation kinetics, with a focus on parameters relevant to living cells.

Though multiple analytical theories treat the looping kinetics of linear chains (16–18), analytical theories for supercoiled DNA have been more challenging to develop. A handful of so-called “repton” models have been proposed to account for the quasi-one-dimensional reptation that is thought to limit loop formation in supercoiled DNA (19–21), yet such models ignore the three-dimensional nature of real DNA and do not distinguish between varying levels of supercoiling. For a more accurate structural representation, many groups have relied on coarse-grained simulation. Monte Carlo simulations based on a discretized WLC model have already established that supercoiling boosts the equilibrium probability of adopting a loopable DNA configuration by roughly two orders of magnitude (22).

Brownian dynamics (BD) simulations have been especially powerful in modeling loop formation kinetics because three-dimensional polymer motions subject to hydrodynamic friction are explicitly represented as a function of time (21,23–27). Using BD to simulate DNA at low ionic strength, Jian et al. report a roughly 100-fold increase in the loop formation rate due to supercoiling, with a slightly larger effect on sites separated by 300 bp than by 600 bp, hinting at a role for supercoiling in shortening the optimal loop length (23). Yet in a subsequent BD study at physiological ionic strength, Huang et al. predict a negligible effect of supercoiling on loop formation rates due to slowed internal motions of a tightly interwound superhelix (24). A pair of BD studies by Klenin et al. also evaluate loop formation timescales in supercoiled DNA although, lacking

a comparison with relaxed DNA, implications for shifts in the length dependence of looping are unclear (25,26). Nonetheless, all BD studies agree that for sites separated by at least 300 bp, loop formation takes place on a time-scale of milliseconds.

Despite their ability to accurately represent polymer dynamics, BD simulations entail extreme computational costs which have limited studies to particular superhelical densities and loop lengths. Notably, none of the aforementioned studies have systematically explored the supercoiling dependence of loop formation kinetics down to the biologically important scale of sub-persistence length loops. On this shorter length scale, and especially in relaxed or weakly supercoiled DNA, bending rigidity drastically suppresses the loop formation rates such that looping events occur rarely over timescales typically accessible with BD.

Here, we present a theoretical model for loop formation kinetics that combines coarse-grained equilibrium structural simulations with a simplistic analytical treatment of polymer dynamics, enabling computationally efficient kinetic predictions across a wide physiological range of superhelical densities and loop lengths. Our treatment of polymer dynamics is an extension of analytical theory introduced by Szabo, Schulten, and Schulten (SSS), in which loop formation is simplified to one-dimensional diffusion along the site-to-site distance reaction coordinate R (17). This approach to dynamics relies on certain assumptions regarding polymer relaxation timescales, and we discuss the range of validity for these assumptions in relation to the parameters we model. Our model starts with Monte Carlo simulations of a discretized WLC representation of supercoiled DNA, leading to looping free energy profiles over R that reflect the full diversity of supercoiled structures. These free energy profiles then define the reaction landscape over which loop formation dynamics are modeled. Our calculations reveal a complex interplay between superhelical density and loop length in controlling loop formation kinetics. We provide structural and energetic interpretations for the trends we observe and discuss their implications for DNA looping in vivo.

MATERIALS AND METHODS

Coarse-grained model of supercoiled DNA

We model a circular DNA molecule as a discrete shearable, stretchable wormlike chain (dssWLC) modified to include a twist-deformation energy and topological constraints, as described previously (28). In this coarse-grained model, the structure of DNA is represented by a string of beads with spatial positions $\vec{r}_i$ and tangent orientation vectors $\vec{u}_i$. The total contour length of the circular chain L_{circle} along with the number of beads m define a discretization length $\Delta = L_{\text{circle}}/m$. The chain conformations are governed by three energetic terms: the conformational free energy of a dssWLC E_{chain} , the quadratic twist-deformation energy E_{twist} , and the steric energy to capture DNA's self-repulsion E_{self} .

The conformational free energy is given by

$$E_{\text{chain}} = \sum_{i=1}^m \left[\frac{\epsilon_b}{2\Delta} \left| \vec{u}_i - \vec{u}_{i-1} - \eta \vec{R}_i^\perp \right|^2 + \frac{\epsilon_\parallel}{2\Delta} \left(\vec{R}_i \cdot \vec{u}_{i-1} - \Delta\gamma \right)^2 + \frac{\epsilon_\perp}{2\Delta} \left| \vec{R}_i \right|^2 \right], \quad (1)$$

where $\vec{R}_i = \vec{r}_i - \vec{r}_{i-1}$ and $\vec{R}_i^\perp = \vec{R}_i - \vec{R}_i \cdot \vec{u}_{i-1}$. Each term in the sum represents one of the three elastic modes captured by the dssWLC—bending, stretching, and shearing—and ϵ_b , $\epsilon_\parallel$, and $\epsilon_\perp$ are their respective coarse-grained moduli. Additionally, $\Delta\gamma$ is the equilibrium contour length of a single segment and η is a bend-shear coupling parameter, describing the tendency of a chain to bend in the direction that it shears. For a chosen discretization length Δ , the dssWLC model (29) provides a procedure to map the desired persistence length l_p (50 nm in this study) onto these coarse-grained elastic parameters. This model exhibits quantitative agreement with the exact results for the WLC model over the range of discretization lengths Δ (29–31), resulting in a model that is adaptable to varying degrees of coarse graining.

We model DNA as having a local twist rate ω , which describes the number of times each single-strand DNA winds around the other per unit contour length. For relaxed DNA, $\omega_0 = (1 \text{ turn})/(10.5 \text{ bp})$. We introduce a quadratic penalty for deformations $\Delta\omega = \omega - \omega_0$ away from the relaxed twist state, given by

$$\frac{E_{\text{twist}}}{k_B T} = \frac{l_t}{2} \oint ds (\Delta\omega)^2 = \frac{(2\pi)^2 (\Delta T w)^2 l_t}{2L}, \quad (2)$$

where l_t is the twist persistence length, s tracks the chain contour, L is the total chain length in nm, and $\Delta T w = (L\Delta\omega)/(2\pi)$ is the total twist deformation of the chain in units of turns. We use $l_t = 70 \text{ nm}$, which agrees with measurements of DNA's torsional rigidity based on topoisomer distributions of small DNA rings generated by ligation (32), but is slightly smaller than the value $l_t \approx 100 \text{ nm}$ measured using a single-molecule rotor bead assay (33). Other methods lead to an even wider range of values, with estimates as low as $l_t \approx 30 \text{ nm}$ and as high as $l_t \approx 120 \text{ nm}$ based on fluorescence polarization anisotropy (34) and cyclization kinetics (35), respectively.

For a given chain conformation, we compute $\Delta T w$ implicitly using the fact that twist and writhe are mathematically linked for a closed circular chain according to $\Delta L k = \Delta T w + W r$. For a particular simulation, $\Delta L k$ is a fixed value denoting the difference in linking number between the simulated chain and a chain with no superhelicity. The writhe $W r$ is a geometric quantity that intuitively represents the number of times the dsDNA chain winds about itself, and is computed for a given chain conformation using Method 1b from Klenin and Langowski (36).

Finally, we include a self-repulsive energy to prevent steric overlap of the chains and to roughly model the effect of varying ionic conditions. Our energy is defined for a given pair of non-neighboring bead-to-bead segments (i, j) , and depends on the distance of closest approach D_{ij} between these segments. For a given conformation, we sum this energy over all pairs of bead-to-bead segments. As in (28), we utilize a standard hard-core Lennard-Jones repulsive interaction, leading to

$$\frac{E_{\text{self}}}{k_B T} = \sum_{i \neq j} \frac{V_{\text{HC}}}{12} \left[\left(\frac{l_{\text{HC}}}{D_{ij}} \right)^{12} - 2 \left(\frac{l_{\text{HC}}}{D_{ij}} \right)^6 + 1 \right] H(l_{\text{HC}} - D_{ij}). \quad (3)$$

The Heaviside step function $H(l_{\text{HC}} - D_{ij})$ dictates that instances where $D_{ij} < l_{\text{HC}}$ contribute to E_{self} , and the contribution is 0 otherwise. The parameters V_{HC} and l_{HC} define the strength and length scale of the interactions, respectively. For all simulations, we use $V_{\text{HC}} = 1$. Simulations in the main text take l_{HC} as the bare diameter of DNA (2 nm), approximating

conditions of high ionic strength where electrostatic repulsion is largely screened. In Fig. S5, we present results using $l_{\text{HC}} = 5 \text{ nm}$, which better describes the conformations of DNA at physiological ionic strength ($\sim 150 \text{ mM}$ monovalent salt) (37). Overall, we reach similar conclusions for both of these simulation conditions, suggesting our conclusions are robust to salt concentrations within the physiological range of conditions.

Monte Carlo simulation procedure

Our Monte Carlo approach for generating supercoiled DNA conformations has been described previously in (28). In brief, we implement a Metropolis algorithm based on two types of structural moves: local chain moves (sub-chain crankshafts and translations) and full chain exchanges between neighboring topological replicas. The amplitudes and windows of the local chain moves are adapted to achieve 50% move acceptance. In addition, we reject any local move that leads to a conformation whose Alexander polynomial differs from 1, indicating the formation of a knot (38).

We initiate a series of parallel simulations (replicas) each defined by a unique value for $\Delta L k$, chosen from a list that evenly spans a chosen range of biologically relevant superhelical densities $\sigma = \Delta L / L k_0 \approx 0.00$ to -0.12 , where $L k_0$ is the linking number of the relaxed plasmid. We note that our model is symmetric with respect to positive and negative supercoiling. Thus, despite the fact that simulations utilized $\sigma \leq 0$, our energetic and kinetic results depend only on the absolute value of σ . Monte Carlo trajectories are generated by performing 10^3 local move trials on each replica, after which neighboring replicas are probabilistically swapped according to a modified Metropolis criterion (39). Each time five replica-exchange trials pass, the conformations are saved. Including replica exchange moves reduces correlations between consecutive saves, especially for the higher $\Delta L k$ ensembles, which are prone to conformational frustration.

The bulk of our study is based on 31 parallel simulations of a 2686 bp (913 nm) plasmid represented by 100 beads (performed for $l_{\text{HC}} = 2 \text{ nm}$ and $l_{\text{HC}} = 5 \text{ nm}$). Each simulation possesses a unique $\Delta L k$ ranging from 0 to -30 , which corresponds to $\sigma \approx 0.00$ to -0.12 for a plasmid of this size. We exclude the most supercoiled replica from our analysis due to its inferior sampling statistics, since this replica experiences half the number of replica-exchange trials relative to replicas with neighbors on both sides. Following a burn-in period containing 10^6 local move trials, simulations are run until 200,000 conformations are saved per $\Delta L k$ replica.

We performed additional sets of simulations in which we either doubled the number of beads or halved the total chain length (or both). We also simulated a 10,000 bp plasmid represented by 300 beads. In all cases, we ran 31 parallel simulations, each defined by a $\Delta L k$ value from a range corresponding to $\sigma \approx 0.00$ to -0.12 . For 200-bead simulations, we saved 50,000 conformations per $\Delta L k$ replica. For the 300-bead simulation, we saved 25,000 conformations per $\Delta L k$ replica. In Figs. S6 and S7, we present a comprehensive view of kinetic results from the $L_{\text{circle}} = 2,686 \text{ bp}$, 200-bead simulation and $L_{\text{circle}} = 10,000 \text{ bp}$, 300-bead simulation, which respectively provide access to shorter and longer loop lengths than the featured simulation in the main text. In Fig. S9, we use these additional simulations to evaluate the effect of discretization length Δ and total chain length L_{circle} on our results.

J-factor calculation

For a given σ and L , the J factor, which expresses the concentration of one site with respect to the other in units of molarity, is computed directly from the simulated distribution of the site-to-site distance R using the formula

$$J_{\sigma L} = \frac{10^{24}}{N_A} \frac{3P_{\sigma L}(R < R_T)}{4\pi R_T^3}, \quad (4)$$

where $P_{\sigma,L}(R < R_T)$ is the number of loops of length L sampled from the σ simulation having $R < R_T$, divided by the total number of loops N_{total} sampled from that simulation. In this study, we use $R_T = 10$ nm.

Looping free energy calculation

To construct free energy profiles, we start by computing binned distributions of site-to-site distance (R) for each superhelical density (σ) and loop length (L) according to

$$P_{\sigma,L}(R_i) = \frac{N_{R_i - \Delta R/2 \leq R < R_i + \Delta R/2}}{N_{\text{total}} \Delta R}, \quad (5)$$

where the numerator is the number of samples with site-to-site distances contained in the bin centered on R_i , N_{total} is the total number of samples, and ΔR is the bin width. Unless otherwise stated, we use 0.04 for ΔR (in units of fractional extension R/L) when constructing free energy profiles for subsequent kinetic analysis, and 0.02 when visualizing free energy profiles, as in Fig. 3. For a given loop length, each saved conformation contains as many measurements of this loop's site-to-site distance as there are beads. Thus, for 100-bead simulations in which 2×10^5 structures are saved, each site-to-site distance distribution contains $N_{\text{total}} = 10^2 \times 2 \times 10^5 = 2 \times 10^7$ samples. We transform these distributions into binned free energy profiles using

$$F_{\sigma,L}(R_i) = -k_B T \log(P_{\sigma,L}(R_i)). \quad (6)$$

These free energy profiles serve as input for the mean first passage time calculations described in the following section.

Mean first passage time for loop formation calculation

The kinetics of loop formation within polymers is a subject of intense theoretical interest, studied both analytically and through simulation. Here, we develop a hybrid approach that extends the SSS analytical theory, in which loop formation is modeled as one-dimensional diffusion across an effective free energy profile along the site-to-site distance reaction coordinate R (17), by providing supercoiling- and loop-length-dependent free energies from coarse-grained simulation. The physical simplicity of the SSS theory leads to a mathematically tractable formula for the mean first passage time for loop formation ($\langle T \rangle$) that depends only on the free energy profile $F_{\sigma,L}(R)$ and an effective diffusion coefficient D , but its usage relies on the assumption of local equilibrium, which states that all internal polymer modes relax on timescales significantly shorter than those associated with loop formation (18,40–42). Though invalid for sufficiently long and flexible polymers due to slow relaxation dynamics, such an assumption is expected to hold for the relatively short, semiflexible chains considered in this study. In the discussion, we draw upon previous theoretical work to support our assumption of local equilibrium, even under conditions of supercoiling.

In this work, we use the following formula for the mean first passage time for loop formation (for complete derivation, see supporting material):

$$\langle T \rangle_{\sigma,L} = \frac{1}{DZ} \int_{R_T}^L dR \int_{R_T}^R dR' \times \int_{R'}^L dR'' e^{-\beta F_{\sigma,L}(R) + \beta F_{\sigma,L}(R') - \beta F_{\sigma,L}(R'')}, \quad (7)$$

where $Z = \int_{R_T}^L e^{-\beta F_{\sigma,L}(R)} dR$ is the unlooped state partition function and R_T

is the site-to-site distance of the transition state, such that $R < R_T$ states are considered “looped” and $R_T < R < L$ states are considered “unlooped.” This formula assumes that the system starts off unlooped and becomes looped upon its first encounter with the transition state.

We have previously used this formalism to model LacI-mediated loop formation kinetics measured using a tethered particle assay (see Eq. S92 in (43) and Eq. 12 in (44)). Our expression for the mean first passage time of loop formation is identical to those from (43) and (44) except that, instead of using exact analytical formulas for the free energies of linear WLCs subject to binding interactions, we generate σ - and L -dependent free energy profiles through exhaustive Monte Carlo simulation of circular dssWLCs, allowing us to map out $\langle T \rangle$ as a function of these two parameters.

We assume that the loop formation time is dominated by the free energy profile, such that D is independent of superhelical density and loop length, and therefore our results are not affected by a particular choice of D . Ambiguity regarding an appropriate parameterization of D is a limitation of our model that prevents rigorous inclusion of solvent-mediated hydrodynamics and the reporting of results in real units of time. Consequently, we report $\langle T \rangle$ in units of the dimensionless time τ_p/D . In the discussion, we consider prior experimental and theoretical work to estimate the impact of ignoring hydrodynamics on our conclusions and to estimate approximate looping time-scales predicted by our model in real units.

For each value of σ and L in our analysis, we numerically compute the triple integral in Eq. 7 using custom code in MATLAB. Instead of using the binned free energy profiles $F_{\sigma,L}(R_i)$, we construct continuous versions $F_{\sigma,L}(R)$ using linear interpolation. For loops larger than roughly a few persistence lengths, states near full extension are entirely absent from our simulations. In these cases, we replace L in Eq. 7 with R_i of the rightmost sample-containing bin.

RESULTS

Overview of the theoretical model

In this work, our goal is to quantify the effect of DNA supercoiling on loop formation rates and to understand the extent to which loop length controls this effect. To this end, we present a model for the DNA loop formation process based on coarse-grained polymer simulation and analytical transition state theory, enabling us to map out the average loop formation time $\langle T \rangle$ (the inverse of the average rate) across a wide range of superhelical densities (σ) and loop lengths (L). Fig. 1 depicts our computational workflow for two particular pairs of σ and L .

First, we simulate equilibrium structural ensembles of a 2686 bp circular DNA plasmid across a physiologically relevant range of superhelical densities ($\sigma \approx 0.000$ to -0.110). As detailed in materials and methods, our simulations model DNA as a discretized WLC using the dssWLC model (29), and we use a replica-exchange Monte Carlo sampling algorithm (28) to achieve highly refined sampling across the full range of parameters we explore.

Next, we compute looping free energy profiles F for each (σ , L) by choosing the site-to-site distance R for a loop of contour length L to be the looping reaction coordinate. States for which $R > R_T$ are considered “unlooped” and states for which $R < R_T$ are considered “looped,” with R_T

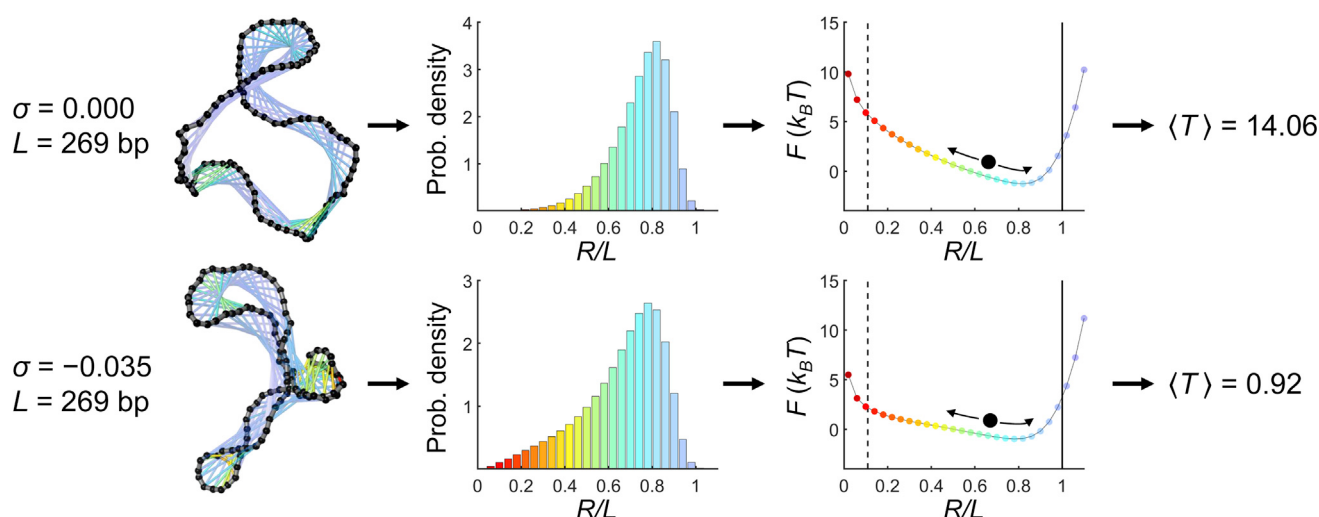


FIGURE 1 Overview of computational approach. First, Monte Carlo simulations are used to generate ensembles of coarse-grained plasmids across a range of superhelical densities σ . For a given (σ , L), distributions of the site-to-site distance R (expressed in fractional units, i.e. R/L , and encoded in color) are tabulated and transformed into free energy profiles. The average loop formation time $\langle T \rangle$ is then computed using Eq. 7, which treats loop formation as a diffusive process along the free energy profile, with a reflecting boundary at $R/L = 1$ (solid vertical line) and an absorbing boundary at the transition state (dashed vertical line).

defining the location of the transition state. We choose $R_T = 10$ nm in this study to reflect the size of a typical prokaryotic looping transcription factor.

Finally, we define the average loop formation time $\langle T \rangle$ as the mean first passage time from unlooped to looped states, which is computed using Eq. 7. Following the approach introduced by SSS (17), this kinetic analysis utilizes a Fokker-Planck equation to model the looping reaction as a diffusive process, and therefore, Eq. 7 considers the entire free energy profile and not just the height and curvature of the transition state, as is done in Kramers' theory (45).

Both the kinetics of looping and the equilibrium probability that DNA adopts a looped configuration may be relevant to regulatory functions in the dynamic context of the cell. We also report DNA looping probabilities based on our simulations (represented as J factors, following common practice in the literature) in Figs. 5, S1, and S10. The rate of formation of a protein-mediated loop is frequently assumed to be proportional to the J factor (6,43). This relation requires that looped DNA configurations are sampled a large number of times before being captured. BD studies predict that looped DNA configurations emerge on a timescale of milliseconds (23–26), yet in vitro studies of LacI-mediated looping dynamics based on tethered particle motion (TPM) report looping on a timescale of tens of seconds (9,46–48), supporting the use of the J factor as a kinetic parameter for this system. The loop-based transcriptional activator NtrC may also function in this regime (49), yet the generality of this assumption across loop lengths, looping systems, and experimental contexts (e.g., in vivo) remains unclear.

In the main text, we focus primarily on the mean first passage time for loop formation $\langle T \rangle$. Overall, $\langle T \rangle$ and J predict

qualitatively similar trends with respect to supercoiling and loop length, although quantitative differences are discussed in the supporting material and in our conclusions.

Coarse-grained structures of supercoiled DNA

Before assessing the full scope of kinetic trends predicted by our model, we focus on particular structural snapshots from our coarse-grained simulations to provide a basic physical intuition for how supercoiling modulates loop formation. In Figs. 2 and S2, we show plasmids at three superhelical densities. Our structures reflect a transition from loose, open rings at $\sigma = 0.000$ toward plectonemic structures that grow progressively tighter as $|\sigma|$ is increased. Each box in Fig. 2 contains two structures at the specified σ , both depicting two sites separated by the specified loop contour length L (134 bp, top; 510 bp, bottom). On the left the sites are in an unlooped configuration, and on the right the sites are looped.

It is immediately clear that plectonemic structures contain many more pairs of sites in close contact than non-supercoiled structures. Indeed, loops of length 134 bp and 510 bp both experience an increase in loop formation rate (a decrease in $\langle T \rangle$) as $|\sigma|$ is increased beyond 0.000.

Yet the precise effect of supercoiling appears to depend strongly on loop length. Between $\sigma = 0.000$ and $\sigma = -0.090$, $\langle T \rangle$ for a 134 bp loop decreases more than 100-fold, yet for a 510 bp loop this factor is less than 5. Furthermore, we observe that supercoiling actually decreases the loop formation rate between $\sigma = -0.035$ and $\sigma = -0.090$ for a 510 bp loop, in stark contrast to a 134 bp loop, which shows an increase in rate with supercoiling across the full range of σ presented in Fig. 2.

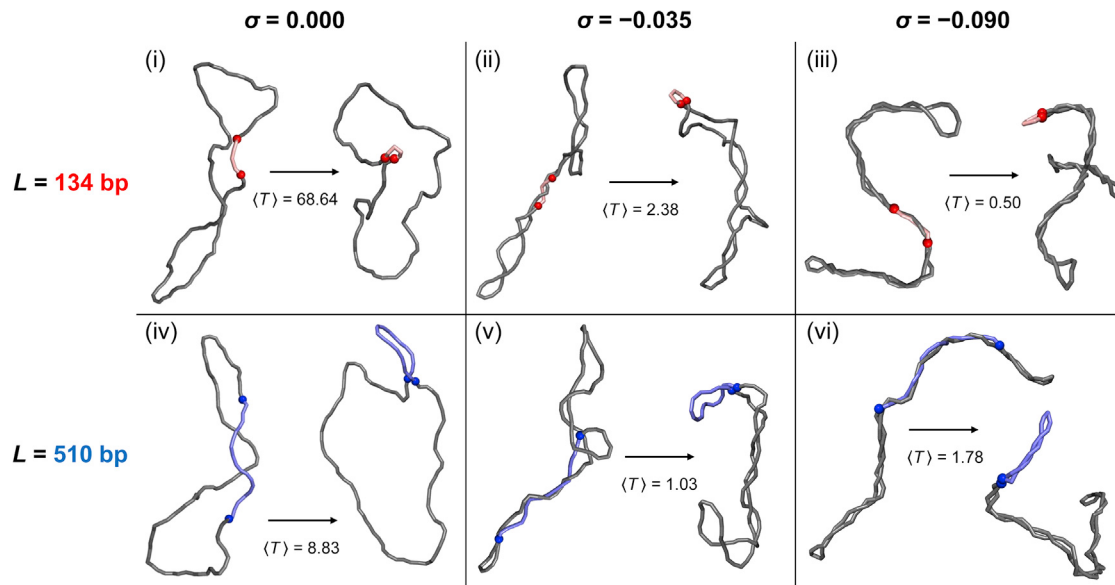


FIGURE 2 Monte Carlo simulation snapshots for six (σ, L) pairs. For each ring structure, colored beads represent the two reactive DNA sites and the intervening colored segment represents a subchain of contour length L . The left and right structures within each box depict unlooped and looped states, respectively. We also report our model prediction $\langle T \rangle$ for each (σ, L) in dimensionless units.

The remainder of this article presents a detailed and comprehensive analysis of these kinetic trends based on the energetics of large-scale structural rearrangements driven by supercoiling. We show that the differential effect of supercoiling originates from an interplay between the loop length and the size scale of plectonemic end loops, on which loops tend to be centered (see looped structures in Fig. 2, boxes (ii), (iii), (v), and (vi)). Because of this positioning preference, loop length controls whether the two DNA sites come together within the end loop region or deeper within the plectonemic stem, and it is the distinct structural responses of these two regions to changes in supercoiling that underlie the kinetic trends predicted by our model.

Looping free energy profiles

While the individual snapshots in Fig. 2 assist in developing a visual interpretation of our results, the looping free energy profiles shown in Fig. 3 depend on thoroughly sampled equilibrium ensembles of simulated structures and, therefore, reveal additional quantitative insights into the physics of loop formation. As in Fig. 1, the vertical black lines define looped and unlooped states for our mean first passage time calculation (Eq. 7), with the dashed line at $R_T = 10$ nm representing an absorbing boundary at the transition state and the solid line at $R = L$ representing a reflecting boundary at full extension. The six curves in Fig. 3 correspond to the six (σ, L) pairs depicted in Fig. 2.

Fig. 3 A shows that the primary effect of supercoiling on a 134 bp loop is to lower the transition state free energy, which

decreases by $\sim 4 k_B T$ between $\sigma = 0.000$ and -0.035 , and by another $\sim 3 k_B T$ between $\sigma = -0.035$ and -0.090 . In contrast, a 510 bp loop only experiences a significant drop in the transition state free energy between $\sigma = 0.000$ and -0.035 (Fig. 3 B), with a comparatively small change in this value between $\sigma = -0.035$ and -0.090 .

These differing energetic trends can be understood in terms of the large-scale structural changes to circular DNA that occur within the two intervals of σ considered. Increasing $|\sigma|$ from 0.000 to 0.035 triggers torsional buckling, leading to the initial emergence of plectonemic structures that feature extended, interwound stem regions capped by bent end loops (compare structures for $\sigma = 0.000$ with those for $\sigma = -0.035$ in Figs. 2 and S2). DNA sites in relaxed plasmids rarely exist in close contact due to the energetic cost of bending the intervening DNA and entropic cost of forming a loop, but in supercoiled plasmids this cost is reduced by the inherent presence of bent end loops, which tend to increase the chance of adopting a looped configuration regardless of length. This effect drives the increase in loop formation between $\sigma = 0.000$ and -0.035 for both loop lengths.

Beyond $|\sigma| = 0.035$, the basic plectoneme has already formed, and increasing $|\sigma|$ further has two primary effects on plasmid structure. First, end loops grow tighter and more severely bent (compare end loops for $\sigma = -0.035$ with those for $\sigma = -0.090$ in Figs. 2 and S2). It is this effect that drives the additional decrease in the transition state free energy (and the additional increase in loop formation rate) between $\sigma = -0.035$ and -0.090 for a 134 bp loop, whose two sites tend to reside at the edge of the tightening

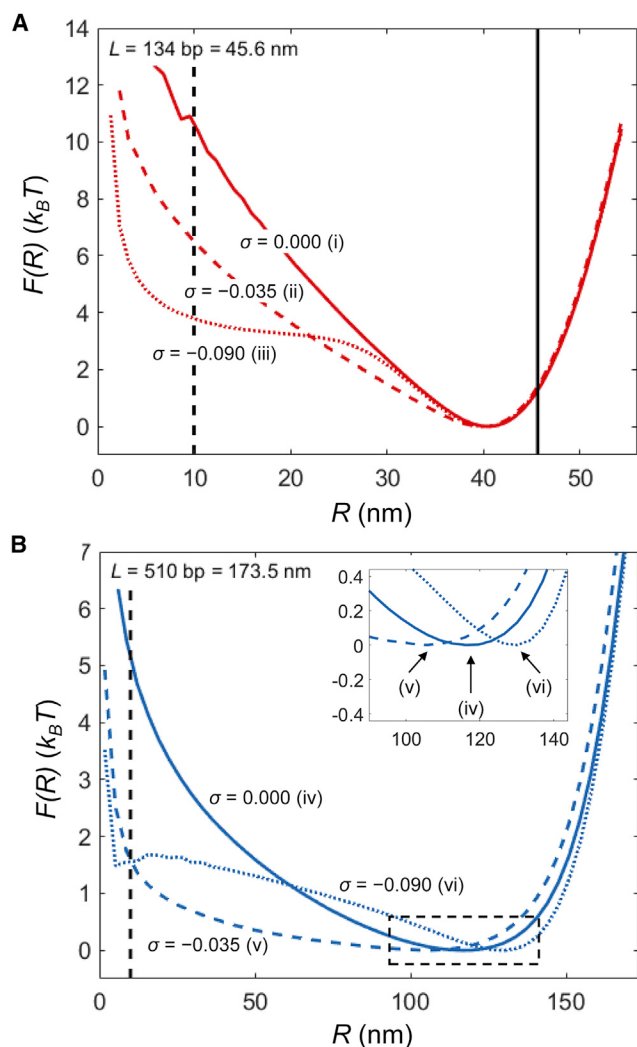


FIGURE 3 Looping free energy profiles. (A) $L = 134$ bp. (B) $L = 510$ bp. The solid, dashed, and dotted curves represent $\sigma = 0.000$, -0.035 , and -0.090 , respectively. The black vertical dashed line denotes the transition state location $R_T = 10$ nm, and the black vertical solid line denotes the fully extended contour length $R = L$. The inset in (B) is a zoom of the boxed region below.

end loop. This structural change has a negligible effect on a 510 bp loop, whose sites reside deeper within the plectonemic stem.

Second, the stem region also tightens, but this is accompanied by a net straightening of chain contours, which facilitates additional writhed crossings that relax twisting energy (again, compare supercoiled structures in Figs. 2 and S2). This straightening manifests as an increase in the free energy of intermediate R states (compare dotted and dashed lines in Fig. 3, A and B), indicating the appearance of additional energy barriers to loop formation at higher $|\sigma|$. This effect is more pronounced for a 510 bp loop than for a 134 bp loop, which already has a tendency to remain straight regardless of σ . Indeed, the free energy minimum for a 510 bp loop shows a considerable shift toward larger R

when σ changes from -0.035 to -0.090 (Fig. 3 B, inset), while the minimum for a 134 bp loop does not.

For a 134 bp loop, the decrease in the transition state free energy going from $\sigma = -0.035$ to -0.090 outweighs the emergence of an energy barrier at intermediate R , so the loop formation rate continues to increase. Yet for a 510 bp loop, this energy barrier is responsible for the increase in $\langle T \rangle$ across the same range of σ .

It is essential to note that for shorter values of L , our simulations contain a non-negligible fraction of apparently overextended loops, with $R > L$ (see Fig. 3 A, curves to the right of the solid black line). This is a known discretization-dependent artifact of the dssWLC model stemming from a quadratic stretch energy that fails to enforce inextensibility. However, a comparison of the $\sigma = 0.000$ energy profile with the energy profile of an ideal linear WLC (Fig. S3 A) for $L = 134$ bp indicates that our simulated energies remain accurate in the range of intermediate to low R . Additionally, we show through numerical means that correcting this artifact near $R \sim L$ has a negligible impact on our kinetic results (Fig. S4). Thus, we relegate the development of an inextensibility correction for the dssWLC model to future work.

Kinetic trends as a function of σ and L

Using 134 bp and 510 bp loops as examples, we have shown that the effect of supercoiling on looping energetics and kinetics depends critically on loop length, with the shorter loop responding to σ more strongly and over a wider range due to its similarity in size to the plectonemic end loop. In this section, we present kinetic results across a fuller range of parameters than previously discussed, providing a comprehensive assessment of whether σ and L may be exploited by living cells to achieve particular gene expression outcomes.

The results in Fig. 4 are based on simulations of a 2686 bp plasmid utilizing $l_{HC} = 2$ nm. Our model predictions are nearly identical with $l_{HC} = 5$ nm, which better approximates physiological ionic conditions, reflecting a weak effect of ionic strength on the coarse-grained conformational properties of supercoiled DNA (Fig. S5). In Fig. S7, we present kinetic results for a 10,000 bp plasmid whose length matches experimental estimates for the size of topological domains within the *Escherichia coli* chromosome (50).

Fig. 4 A shows how $\langle T \rangle$ responds to supercoiling for eight particular loop lengths. Although we have already described extreme quantitative differences in this response between a 134 bp and a 510 bp loop, the curves shown here reveal a qualitative universality across loop lengths, reflecting common underlying physics. Namely, as $|\sigma|$ increases, all loop lengths initially exhibit a reduction in the loop formation time toward a local minimum for $\langle T \rangle$, at which point further increases to $|\sigma|$ cause $\langle T \rangle$ to increase. We see that loop length determines both the overall magnitude of these

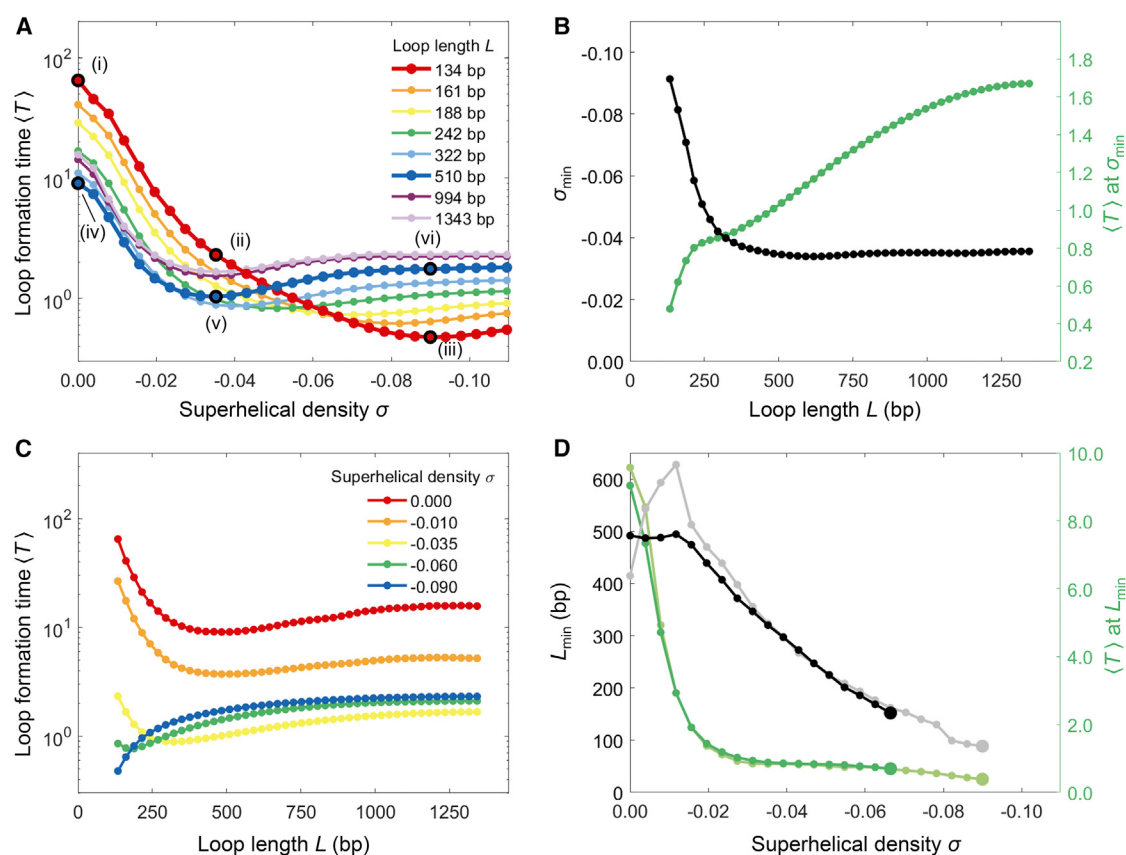


FIGURE 4 Loop formation time $\langle T \rangle$ versus superhelical density σ and loop length L . (A) $\langle T \rangle$ versus σ curves for particular values of L . The annotated data points correspond to the (σ, L) pairs analyzed in Figs 2 and 3. (B) Location of the local minima in (A) as a function of L . (C) $\langle T \rangle$ versus L curves for particular σ values. (D) Location of the local minima in $\langle T \rangle$ versus L curves as a function of σ for 100-bead (darker curves) and 200-bead (lighter curves) simulations. Sample $\langle T \rangle$ versus L curves for the 200-bead simulation are shown in Fig. S6 B. For all curves in (D), the large data point toward the right denotes the largest $|\sigma|$ for which we observe a local minimum. In (B) and (D), minima are determined by fitting a parabola to the minimum data point and its two nearest neighbors.

changes and the superhelical density $\sigma_{\min}$ for which $\langle T \rangle$ is minimized (Fig. 4 B).

For all loop lengths, the non-monotonic kinetic response stems directly from the two competing structural effects described in the previous section: supercoiling-driven site concentration, which reduces loop formation time by lowering the transition state free energy, and supercoiling-driven chain straightening, which decreases loop formation rate by suppressing intermediate R states along the reaction pathway. For a particular loop length in Fig. 4 A, the location of the minimum is the value of σ at which these offsetting effects cancel. We point out that in an earlier modeling study, these specific large-scale structural effects (torsional buckling followed by chain straightening) were also shown to underlie a non-monotonic response of a plasmid's hydrodynamic radius to supercoiling (28).

Fig. 4 B shows that for loops longer than $\sim 2l_p$, $\sigma_{\min}$ is independent of loop length and roughly equal to -0.035 , in accord with the notion that loops in this length range experience site concentration primarily during the initial buckling transition. Shorter loops, however, experience site

concentration over wider ranges of σ due to tightening of the end loops, and therefore, $\sigma_{\min}$ takes on higher absolute values.

A number of in vitro and in vivo biophysical studies specifically measure the length dependence of looping. To compare our model predictions with these data, Fig. 4 C presents an alternative view of the data set shown in Fig. 4 A, with the abscissa and legend variables swapped to now emphasize loop length trends at constant σ .

These curves also tend to have local minima, with the location of these minima (plotted in Fig. 4 D) representing the length of the most readily forming loop at a particular σ . For $\sigma = 0.000$, $L_{\min}$ equals 492 bp, which closely agrees with our prediction for $L_{\min}$ of simulated linear chains (see Fig. S8), reflecting similarities in the conformational properties of linear chains and those embedded within relaxed plasmids. Initially, increasing $|\sigma|$ primarily shifts $\langle T \rangle$ versus L curves downward, without having a major effect on $L_{\min}$. Then, beyond the initial buckling transition ($|\sigma| \geq 0.015$), $L_{\min}$ decreases nearly linearly with σ . Past $|\sigma| \approx 0.066$, $\langle T \rangle$ versus L curves for the 100-bead simulation no longer

show local minima, presumably because these minima occur below 134 bp, which is the shortest loop we analyze for this discretization length. However, $\langle T \rangle$ versus L curves for the 200-bead simulation allow us to resolve local minima for slightly higher values of $|\sigma|$ (see Fig. S6 B, which leads to the pale curves in Fig. 4 D), confirming that the linear relationship continues toward shorter loop lengths. This trend has major implications for DNA looping in vivo, where chromosomes are known to be supercoiled.

DISCUSSION

The length dependence of looping in vitro and in vivo

Our model reveals a clear mechanism by which supercoiling shifts the optimal loop length toward shorter values—namely, shorter loops display an enhanced acceleration due to supercoiling because of the proximity of their reactive sites to the plectonemic end loop. In Figs. 5 and S10, we directly compare our predictions for the loop length dependence of the average looping rate $\langle T \rangle^{-1}$ with various measurements of DNA looping efficiency as a function of loop length both in vitro and in vivo. Our goal is not to suggest that our model fully captures the in vivo mechanical environment of DNA, which includes architectural proteins (51–59) and non-Brownian ATP-driven fluctuations (60,61). Moreover, BD simulations have revealed that an increased loop capture radius can shorten the optimal loop length,

demonstrating that multiple distinct mechanisms may contribute to these effects (41). Instead, we wish to simply explore when the isolated effects of supercoiling on the length dependence of looping are sufficient to explain experimental trends.

To best illustrate shifts in the loop length dependence across a diverse set of assays and loop formation metrics, all data sets and model predictions in Figs. 5 and S10 are normalized to their respective maxima. We additionally plot J versus L predictions from the same simulations as our $\langle T \rangle^{-1}$ predictions. Recall that $\langle T \rangle^{-1}$ determines the rate of stable loop formation when a stable loop forms the first time the DNA visits a looped configuration, while J controls this rate when looped DNA configurations are visited numerous times before a stable loop is formed. Thus, depending on which of these kinetic scenarios better describes a particular looping reaction, either $\langle T \rangle^{-1}$ or J may accurately capture trends in looping efficiency.

In Fig. 5 A, we observe a close correspondence between the in vitro data, which are acquired under conditions of no supercoiling, and our predictions at $\sigma = 0.000$ for both $\langle T \rangle^{-1}$ and J , with clear peaks in the vicinity of 500 bp. We display results from our $L_{\text{circle}} = 10,000$ bp simulations in Fig. 5 to enable experimental comparisons across a wider range of loop lengths. Due to the larger discretization length in this simulation relative to ones with shorter rings, our resolution in determining the location of the peak loop length is diminished, especially when this length is shortened due to supercoiling. However,

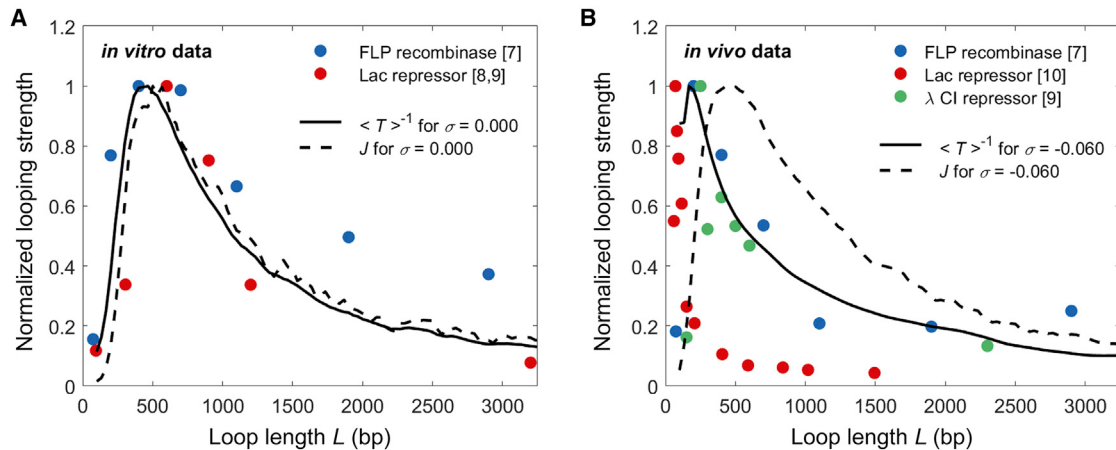


FIGURE 5 Comparison of loop length dependence predictions with experimental looping data. (A) $\langle T \rangle^{-1}$ versus L and J versus L predictions from $L_{\text{circle}} = 10,000$ bp, $\sigma = 0.000$ simulations (solid and dashed black curves) compared with in vitro looping data (colored circles). The FLP recombinase series of data (blue circles) are recombination percentages from Fig. 2 D ([FLP] = 12 nM data) of (7). The Lac repressor series of data (red circles) was generated by associating the average TPM J factors from Fig. 10, A and B of (8) with the average lengths plotted in each figure (94 and 305 bp, respectively), and combining these data with the TPM J -factor data in Fig. 5 E of (9) (600, 900, 1200, and 3200 bp). (B) $\langle T \rangle^{-1}$ versus L and J versus L predictions from $L_{\text{circle}} = 10,000$ bp, $\sigma = -0.060$ simulations (solid and dashed black curves) compared with in vivo looping data (colored circles). The FLP recombinase series of data (blue circles) are reporter/reference recombination ratios taken from Fig. 7 C of (7). The Lac repressor series of data (red circles) are repression ratios from Fig. 3 B of (10). The λ CI series of data (green circles) are in vivo J factors from Fig. 4 of (9). We note that (9) also reports in vivo J factors for Lac-mediated loops with lengths no shorter than 242 bp, but we have omitted them from this figure because these data fail to display a maximum value. For all curves or series of data, values are normalized to the maximum to facilitate comparisons of looping peak locations.

we show in Fig. S10 that for both $\langle T \rangle^{-1}$ and J , the location of the peak loop length is only weakly sensitive to L_{circle} .

In contrast to the in vitro looping data, Fig. 5 B shows that the in vivo looping data, which peak at shorter loop lengths, are better described by our $\langle T \rangle^{-1}$ predictions for $\sigma = -0.060$, a value we have chosen to approximate superhelical densities in living cells. The closer match between these data and our model's predictions for $\sigma = -0.060$ relative to $\sigma = 0.000$ —especially for the assays based on FLP recombination in human cells and λ CI-mediated repression in *E. coli* cells—supports the notion that the effect of supercoiling on loop formation rates at least partially explains the shifts in loop length dependence in vitro versus in vivo. However, we emphasize that this correlation does not rule out alternative explanations for these shifts. The fact that the LacI-based assay shows a different peak (~ 71 bp) versus the λ CI-based assay (~ 250 bp), despite both occurring within *E. coli*, reveals a limitation of our model, and may reflect constraints due to protein geometry (62–64) or a role for architectural proteins such as HU, HMGB, and Fis (51–59,64).

Interestingly, our analysis reveals that although $\langle T \rangle^{-1}$ versus L curves experience a significant shift toward shorter optimal loop lengths due to supercoiling, J versus L curves do not (compare dashed curves between Fig. 5, A and B). This observation suggests that for looping systems that favor shorter loop lengths in vivo, regulatory function may be limited by the rate at which DNA can adopt a looped configuration rather than only the equilibrium looping probability.

Exact comparisons based on in vivo estimates for σ are challenging because in vivo supercoiling is organism dependent, highly dynamic, and partially constrained through stable protein-DNA interactions (65). Measurements on plasmids extracted from *E. coli* show σ ranging from -0.040 to -0.080 (66–70), depending on growth conditions; about half of the supercoils in *E. coli* may be unconstrained (71,72). In one study, reported estimates for the unconstrained superhelical density generated by transcription in eukaryotic cells were in the range of -0.060 to -0.070 (73). Considering a wide range of potentially physiologically relevant superhelical densities from -0.020 to -0.080 , Fig. 4 D predicts peak loop lengths from ~ 100 to a few hundred basepairs.

Supercoiling as a kinetic control knob

Through evolution over many generations, genomes utilizing loop-based regulation can adjust the distance between looping sites to achieve particular regulatory outcomes. However, within a single generation loop length is not adjustable, and it is more likely that cells would utilize supercoiling to modulate looping over these shorter time-scales. Supercoiling is highly dynamic in vivo, with a wide range of phenomena having large transient effects on in vivo superhelical densities, including transcription,

nucleosome assembly and disruption, and topoisomerase activity. Indeed, supercoiling has been described as a “second messenger,” transmitting information about environmental cues such as osmolarity, nutrient availability, pH, and temperature to genetic regulatory networks within the cell (74,75).

Our model allows us to quantitatively assess how dynamic changes in supercoiling affect the looping process. All loop lengths depicted in Fig. 4 A display a significant increase in loop formation rate as the initial supercoiled structure is formed (i.e., between $\sigma = 0.000$ and -0.015). Starting from a relaxed state, even a modest pulse of supercoiling is predicted to increase the formation rate of short loops by more than an order of magnitude, while for longer loops this increase caps out at around fivefold.

Yet, for higher values of $|\sigma|$ —while still remaining within the physiologically relevant range—the supercoiling response tends to flatten out. Though changing σ from -0.035 to -0.090 has an inverted effect on short versus long loops, none of the loop lengths plotted in Fig. 4 A experience larger than a fivefold change in $\langle T \rangle$ across this interval, with the modulation not even reaching twofold for loops longer than ~ 200 bp. Thus, it appears that the mechanics of DNA bending and twisting may actually serve to buffer large effects of supercoiling on loop formation at intermediate superhelical densities for all but the shortest of loops.

Although we wish to understand supercoiling's impact on looping in the context of a living genome, in vitro experiments on artificial but supercoiled DNA substrates lack additional complexities encountered in vivo, and may provide better tests of our model. In vitro transcription assays with the glnAp2-NtrC promoter-enhancer on 7649 bp plasmids reveal that supercoiling ($\sigma \approx -0.070$) enhances formation of 110 bp and 2500 bp loops roughly 2-fold and 50-fold, respectively (76). Although our model captures the supercoiling-induced acceleration for both loop lengths, it also predicts a significantly larger acceleration for 110 bp than the enhancement observed experimentally. For this system, the protein complex that bridges the enhancer and promoter contains multiple NtrC activators, the kinase NtrB, σ^{54} , and RNA polymerase and may approach the size of a 110 bp DNA loop itself. Thus, sliding- or oligomerization-based mechanisms may facilitate high transcriptional activation even in the absence of supercoiling, potentially explaining deviations from our theory for this loop length.

In a follow-up study, Polikanov et al. analyze the 2500 bp loop using BD simulations, concluding that the J factor better accounts for the magnitude of the observed supercoiling response than the looping rate (49). Our J factor and rate calculations are in qualitative agreement with these BD results (including a non-monotonicity in the rate's response to supercoiling), lending support to their conclusion. Our model provides an additional test: if the glnAp2-NtrC system does indeed operate in kinetic regime governed by J , then the optimal loop length will be roughly 500 bp, with

a negligible dependence on supercoiling (compare *dashed curves* in Fig. 5, A and B).

Though not an exact comparison because of linearized and extended DNA substrates, in vitro magnetic manipulation experiments also support our prediction that supercoiling accelerates loop formation (46,77,78). However, these studies have not explored the loop length dependence of this effect, so their ability to challenge our model remains limited.

The supercoiling dependence of LacI-mediated looping in *E. coli* has been measured using growth phase and topoisomerase mutants to probe different in vivo superhelical densities. Studying the natural operon, which supports three loops of lengths 92 bp, 401 bp, and 493 bp, Fulcrand et al. report a threefold decrease in β -galactosidase repression (a measure of looping) between early and late exponential phases, accompanied by a decrease in superhelical density from $\sigma \approx -0.076$ to $\sigma \approx -0.062$ (70). Using the natural operon and a series of single loop operators with lengths near 80 bp, Mogil et al. observe a similar trend, reporting two- to threefold increases and decreases in β -galactosidase repression for, respectively, topoisomerase-deficient ($\sigma \approx -0.07$) and gyrase-deficient ($\sigma \approx -0.04$) mutants relative to the wild-type strain ($\sigma \approx -0.05$) (69). Together, these data indicate a slight repressive effect of additional negative supercoiling within the physiological range, which accords with our prediction for sub-persistence length loops, although the presence of longer loops in the natural operon complicates an exact comparison.

For loop lengths larger than ~ 500 bp, Priest et al. report that looping J factors for LacI are roughly 10- to 20-fold larger in vivo than in vitro, without a significant dependence on loop length (9). Our predictions for $\langle T \rangle$ and J capture this comparison as long as we assume that the in vivo superhelical densities exceed $|\sigma| \approx 0.035$. If we compare our model's $\langle T \rangle$ versus L curve for $\sigma = 0.000$ (*red curves* in Figs. 4 C, S6 B, and S7 B) with the cluster of curves at higher $|\sigma|$ (*bluer curves* in Figs. 4 C, S6 B, and S7 B), we see that supercoiling reduces loop formation time roughly 10-fold, with almost no dependence on either loop length or superhelical density. J factors also show a minimal dependence on loop length and superhelical density in this range, although the predicted enhancement due to supercoiling is closer to 100-fold (compare *red* with *bluer curves* in Fig. S1 B).

Assumptions underlying the SSS approach

The SSS treatment of loop formation as diffusion along the single reaction coordinate R relies on an assumption of local equilibrium that internal polymer modes relax on timescales shorter than those associated with loop formation. For polymer chains comprising many flexible elements, slow relaxation dynamics invalidate this assumption, requiring use of the more complex theory of Wilemski and Fixman (WF),

which solves the three-dimensional diffusion equation in the presence of sink functions that promote site-to-site contact (16,79). Yet semiflexible polymers such as dsDNA are rigid on length scales associated with loop capture ($L_{\text{Kuhn}} > R_T$), and as long as the total number of effective Kuhn lengths is not exceedingly large, local equilibrium is expected to hold (18,40–42). In support of this notion, BD simulations of linear WLCs comprising an intermediate number of effective Kuhn lengths ($\sim 1 - 5$) exhibit the $\langle T \rangle \sim L^{3/2}$ scaling predicted by SSS (41), in contrast to the $\langle T \rangle \sim L^2$ scaling predicted by WF (16,79).

Our model predicts shorter loop formation timescales in supercoiled WLCs, calling into question whether local equilibrium persists under such conditions. Unlike for linear WLCs, we lack analytical results that can be compared with BD simulations to test for local equilibrium, but simulations of supercoiled WLCs still allow us to explicitly compare relaxation timescales associated with relevant configurational modes. For moderately supercoiled DNA with ring lengths in the 3 kbp range, loop formation occurs on a timescale of 1–10 ms (23–26). Writhe relaxes on a timescale of 1–10 μ s (23,80,81), radius of gyration on a timescale of 10–100 μ s (23,24,81), and the number of superhelix branches on a timescale of 1 ms (24). Thus, branching appears to be the only process with a rate as slow as loop formation, but for rings shorter than 3 kbp, roughly 80% of structures are unbranched (25), and collisions between sites on separate branches account for at most 20% of looping events (24). For longer ring lengths, including the 10 kbp ring analyzed in Fig. S7, slow branch relaxation is more likely to cause violation of local equilibrium under conditions of supercoiling.

A major limitation of our kinetic model is its reliance on a single parameter D to describe diffusive motion along R for all loop lengths and superhelical densities in our analysis. The three-dimensional nature of polymers obscures an exact physical interpretation of D , but it should roughly capture hydrodynamic effects that limit timescales of motion through the solvent. Basic hydrodynamics necessitate that D decreases with loop length L , although single-molecule measurements of DNA's translational diffusion coefficient (not an exact proxy for internal motions along R) reveal a relatively weak dependence $D_{\text{trans}} \sim L^{-0.58}$, with a negligible dependence on topological state. Similar considerations lead us to conclude that D ought to depend weakly on R due to the coupling between this reaction coordinate and the relative site-to-site orientation. The dependence of D on superhelical density σ is particularly difficult to judge, with studies providing varying evidence on this matter. Certain BD simulations report a slowing of internal dynamics upon supercoiling (24) while fluorescence correlation measurements of internally labeled plasmids report an acceleration (82,83). By assuming that D is independent of L and σ , our calculations specifically model the effect of L - and σ -dependent conformational energetics on loop

formation kinetics. As these matters become clearer, it will be straightforward to modify our approach based on L - and σ -dependent diffusion coefficients.

Despite ambiguity in parameterizing D , it is easier to believe that a one-dimensional model along R may accurately capture dynamics for supercoiled DNA, which assumes a quasi-one-dimensional plectonemic structure, than for linear DNA. Simulating a 1.5 kbp plasmid at $|\sigma| \approx 0.056$ using BD, Wedemann et al. estimate that the contour position of the plectonemic end loop diffuses according to $D \approx 2.5 \times 10^6 \text{ nm}^2/\text{s}$ (84). Roughly interpreting this value as an estimate for D (and using $l_p = 50 \text{ nm}$), our model predicts that most loops form on a timescale of milliseconds, consistent with direct estimates from BD simulation (23–26,85).

Additional model simplifications

Our model does not address the potential effect of supercoiling on loop dissociation rates, which could also be relevant in vivo. Dissociation rates of protein-mediated loops are often assumed to be independent of loop geometry and external stress with the justification that molecular dissociation rates are local phenomena defined primarily by properties of the intermolecular interface. In support of this view, BD simulations of a 3 kbp plasmid containing two mutually attractive sites situated on opposite sides of the ring predict that supercoiling primarily increases loop formation rates by facilitating rapid rebinding, having virtually no effect on dissociation rates (27). Yet, for short ($\sim 100 \text{ bp}$) DNA loops stabilized by LacI, experimental dissociation rates were found to decrease with the J factor (43), indicating that mechanical stresses within the DNA loop can drive loop breakdown. Whether this holds true for longer lengths or other looping systems remains unclear.

Finally, our model makes no distinction between positive and negative superhelicity, even though it is well established that DNA has differing mechanical responses to these two types of forces (33). In particular, negative torsional stress drives the sequence-dependent formation of alternative DNA conformations such as Z-DNA, strand-separated DNA, and cruciform DNA (86–88). At a given superhelical density, these transitions will compete with plectoneme buckling and extension, buffering the effect of superhelicity on the unconstrained supercoiling that facilitates looping within our model. As a result, higher negative superhelical densities may be needed to achieve the same effect on looping as a given positive superhelical density. To further complicate the matter, alternative forms of DNA may lead to localized defects in structural or mechanical properties (e.g., kinks or soft spots), with significant consequences on loop formation (89,90). Incorporating these additional complexities into future models will be valuable to our overall understanding of looping phenomena.

CONCLUSION

In this article, we present a detailed analysis of the effect of supercoiling on DNA looping kinetics, with the intention of elucidating the extent to which this effect drives observed shifts in looping behavior in vitro versus in vivo. Enabled by a methodology that permits us to predict kinetics over a broad range of biophysical variables, we report a rich interplay between superhelical density and loop length in determining loop formation rates, and we interpret this interplay on the basis of large-scale conformational changes in supercoiled DNA. We show that physiological levels of supercoiling do in fact contribute to the apparent preference for shorter loops in vivo. We emphasize that our model makes no reference to any particular organism or looping system, and thus the insights we provide are relevant in a wide range of scenarios. On the other hand, it is clear that system-specific details such as protein geometry need to be considered in future models in order to achieve precise quantitative agreement with experimental data. Finally, our model predicts a non-monotonic kinetic response to supercoiling that is tuned by loop length, enabling us to quantify some limits of a cell's ability to modulate looping behavior through dynamic supercoiling.

To summarize, our work provides quantitative physical insight into existing measurements of DNA looping, reveals physical principles that will inform future efforts to both study and utilize the link between supercoiling and loop formation, and establishes a computational framework that can be augmented to capture additional system-specific complexities pertinent to the looping process.

AUTHOR CONTRIBUTIONS

C.H.S. designed research, performed research, contributed analytic tools, analyzed data, and wrote the article; Z.B. designed research; A. J. S. designed research and wrote the article.

SUPPORTING MATERIAL

Supporting material can be found online at <https://doi.org/10.1016/j.bpj.2022.04.009>.

ACKNOWLEDGMENTS

Financial support for this work is provided by the National Science Foundation, Physics of Living Systems Program (PHY-1707751).

REFERENCES

1. Vilar, J., and L. Saiz. 2005. DNA looping in gene regulation: from the assembly of macromolecular complexes to the control of transcriptional noise. *Curr. Opin. Genet. Dev.* 15:136–144.
2. Bouwman, B., and W. de Laat. 2015. Getting the genome in shape: the formation of loops, domains, and compartments. *Genome Biol.* 16:154.

3. Schleif, R. 1992. DNA looping. *Annu. Rev. Biochem.* 61:199–223.
4. Shimada, J., and H. Yamakawa. 1984. Ring-closure probabilities for twisted wormlike chains. Applications to DNA. *Macromolecules.* 17:689–698.
5. Spakowitz, A. 2006. Wormlike chain statistics with twist and fixed ends. *Europhys. Lett.* 73:684–690.
6. Shore, D., and R. Baldwin. 1983. Energetics of DNA twisting. I. Relation between twist and cyclization probability. *J. Mol. Biol.* 170:957–981.
7. Ringrose, L., S. Chabanis, ..., A. F. Stewart. 1999. Quantitative comparison of DNA looping in vitro and in vivo: chromatin increases effective DNA flexibility at short distances. *EMBO J.* 18:6630–6641.
8. Han, L., H. Garcia, ..., R. Phillips. 2009. Concentration and length dependence of DNA looping in transcriptional regulation. *PLoS One.* 4:e5621.
9. Priest, D. G., L. Cui, ..., K. E. Shearwin. 2014. Quantitation of the DNA tethering effect in long-range DNA looping in vivo and in vitro using the Lac and λ repressors. *Proc. Natl. Acad. Sci. U S A.* 111:349–354.
10. Müller, J., S. Oehler, and B. Müller-Hill. 1996. Repression of lac promoter as a function of distance, phase and quality of an auxiliary lac operator. *J. Mol. Biol.* 257:21–29.
11. Brunwasser-Meirom, M., Y. Pollak, ..., R. Amit. 2016. Using synthetic bacterial enhancers to reveal a looping-based mechanism for quenching-like repression. *Nat. Commun.* 7:1–12.
12. Becker, N. A., T. L. Schwab, ..., L. J. Maher, III. 2018. Bacterial gene control by DNA looping using engineered dimeric transcription activator like effector (TALE) proteins. *Nucleic Acids Res.* 46:2690–2696.
13. Hao, N., K. Shearwin, and I. Dodd. 2017. Programmable DNA looping using engineered bivalent dCas9 complexes. *Nat. Commun.* 8:1628.
14. Hao, N., K. Shearwin, and I. Dodd. 2019. Positive and negative control of enhancer-promoter interactions by other DNA loops generates specificity and tunability. *Cell Rep.* 26:2419–2433.
15. Tse, D., N. Becker, ..., L. Maher. 2021. Designed architectural proteins that tune DNA looping in bacteria. *Nucleic Acids Res.* 49:10382–10396.
16. Wilemski, G., and M. Fixman. 1974. Diffusion-controlled intrachain reactions of polymers. I Theory. *J. Chem. Phys.* 60:866–877.
17. Szabo, A., K. Schulten, and Z. Schulten. 1980. First passage time approach to diffusion controlled reactions. *J. Chem. Phys.* 72:4350–4357.
18. Toan, N., G. Morrison, ..., D. Thirumalai. 2008. Kinetics of loop formation in polymer chains. *J. Phys. Chem. B.* 112:6094–6106.
19. Sessions, R., M. Oram, ..., S. Halford. 1997. Random walk models for DNA synapsis by resolvase. *J. Mol. Biol.* 270:413–425.
20. Marko, J. 1997. The internal 'slithering' dynamics of supercoiled DNA. *Physica A.* 244:263–277.
21. Klenin, K., and J. Langowski. 2001. Kinetics of intrachain reactions of supercoiled DNA: theory and numerical modeling. *J. Chem. Phys.* 114:5049–5060.
22. Vologodskii, A. V., S. D. Levene, ..., N. R. Cozzarelli. 1992. Conformational and thermodynamic properties of supercoiled DNA. *J. Mol. Biol.* 227:1224–1243.
23. Jian, H., T. Schlick, and A. Vologodskii. 1998. Internal motion of supercoiled DNA: brownian dynamics simulations of site juxtaposition. *J. Mol. Biol.* 284:287–296.
24. Huang, J., T. Schlick, and A. Vologodskii. 2001. Dynamics of site juxtaposition in DNA. *Proc. Natl. Acad. Sci. U S A.* 98:968–973.
25. Klenin, K., and J. Langowski. 2001. Diffusion-controlled intrachain reactions of supercoiled DNA: Brownian dynamics simulations. *Biophysical J.* 80:69–74.
26. Klenin, K., and J. Langowski. 2001. Intrachain reactions of supercoiled DNA simulated by Brownian dynamics. *Biophysical J.* 81:1924–1929.
27. Benedetti, F., J. Dorier, and A. Stasiak. 2014. Effects of supercoiling on enhancer-promoter contacts. *Nucleic Acids Res.* 42:10425–10432.
28. Krajina, B. A., and A. J. Spakowitz. 2016. Large-scale conformational transitions in supercoiled DNA revealed by coarse-grained simulation. *Biophysical J.* 111:1339–1349.
29. Koslover, E., and A. Spakowitz. 2013. Discretizing elastic chains for coarse-grained polymer models. *Soft Matter.* 9:7016–7027.
30. Spakowitz, A. J., and Z.-G. Wang. 2004. Exact results for a semiflexible polymer chain in an aligning field. *Macromolecules.* 37:5814–5823.
31. Spakowitz, A. J., and Z.-G. Wang. 2005. End-to-end distance vector distribution with fixed end orientations for the wormlike chain model. *Phys. Rev. E.* 72:041802.
32. Frank-Kamenetskii, M., A. Lukashin, ..., A. Vologodskii. 1985. Torsional and bending rigidity of the double helix from data on small DNA rings. *J. Biomol. Struct. Dyn.* 2:1005–1012.
33. Bryant, Z., M. Stone, ..., C. Bustamante. 2003. Structural transitions and elasticity from torque measurements on DNA. *Nature.* 424:338–341.
34. Millar, D., R. Robbins, and A. Zewail. 1980. Direct observation of the torsional dynamics of DNA and RNA by picosecond spectroscopy. *Proc. Natl. Acad. Sci. U S A.* 77:5593–5597.
35. Crothers, D., J. Drak, ..., S. Levene. 1992. DNA bending flexibility, and helical repeat by cyclization kinetics. *Methods Enzymol.* 212:3–29.
36. Klenin, K., and J. Langowski. 2000. Computation of writhe in modeling of supercoiled DNA. *Biopolymers.* 54:207–317.
37. Rybenkov, V. V., N. R. Cozzarelli, and A. V. Vologodskii. 1993. Probability of DNA knotting and the effective diameter of the DNA double helix. *Proc. Natl. Acad. Sci. U S A.* 90:5307–5311.
38. Frank-Kamenetskii, M., and V. Vologodskii. 1981. Topological aspects of the physics of polymers: the theory and its biophysical applications. *Uspekhi Fizicheskikh Nauk.* 24:679–695.
39. Hukushima, K., and K. Nemoto. 1996. Exchange Monte Carlo method and application to spin glass simulations. *J. Phys. Soc. Jpn.* 65:1604–1608.
40. Jun, S., J. Bechhoefer, and B.-Y. Ha. 2003. Diffusion-limited loop formation of semiflexible polymers: Kramers theory and the intertwined time scales of chain relaxation and closing. *Europhysics Lett.* 64:420–426.
41. Chen, J., H. Tsao, and Y. Sheng. 2004. First-passage time of cyclization dynamics of a wormlike polymer. *Europhysics Lett.* 65:407–413.
42. Chen, J., H. Tsao, and Y. Sheng. 2005. Diffusion-controlled first contact of the ends of a polymer: crossover between two scaling regimes. *Phys. Rev. E.* 72:031804.
43. Chen, Y.-J., S. Johnson, ..., R. Phillips. 2014. Modulation of DNA loop lifetimes by the free energy of loop formation. *Proc. Natl. Acad. Sci. U S A.* 111:17396–17401.
44. Mulligan, P., Y.-J. Chen, ..., A. Spakowitz. 2015. Interplay of protein binding interactions, DNA mechanics, and entropy in DNA looping kinetics. *Biophysical J.* 109:618–629.
45. Kramers, H. 1940. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica.* 7:284–304.
46. Normanno, D., F. Vanzi, and F. S. Pavone. 2008. Single-molecule manipulation reveals supercoiling-dependent modulation of lac repressor-mediated DNA looping. *Nucleic Acids Res.* 36:2505–2513.
47. Milstein, J., Y. Chen, and J. Meiners. 2010. Bead size effects on protein-mediated DNA looping in tethered-particle motion experiments. *Biopolymers.* 95:144–150.
48. Johnson, S., M. Lindén, and R. Phillips. 2012. Sequence dependence of transcription factor-mediated DNA looping. *Nucleic Acids Res.* 40:7728–7738.
49. Polikanov, Y., V. Bondarenko, ..., V. Studitsky. 2007. Probability of the site juxtaposition determines the rate of protein-mediated DNA looping. *Biophysical J.* 93:2726–2731.
50. Postow, L., C. Hardy, ..., N. Cozzarelli. 2004. Topological domain structure of the *Escherichia coli* chromosome. *Genes Development.* 18:1766–1779.

51. Becker, N., J. Kahn, and L. Maher. 2005. Bacterial repression loop require enhanced DNA flexibility. *J. Mol. Biol.* 349:716–730.
52. Becker, N., J. Kahn, and L. Maher. 2007. Effects of nucleoid proteins on DNA repression loop formation in *Escherichia coli*. *Nucleic Acids Res.* 35:3988–4000.
53. Becker, N., J. Kahn, and L. Maher. 2008. Eukaryotic HMGB proteins as replacements for HU in *E. coli* repression loop formation. *Nucleic Acids Res.* 36:4009–4021.
54. Becker, N., and L. Maher. 2015. High-resolution mapping of architectural DNA binding protein facilitation of a DNA repression loop in *Escherichia coli*. *Proc. Natl. Acad. Sci. U S A.* 112:7177–7182.
55. Bond, L., J. Peters, ..., L. Maher. 2010. Gene repression by minimal *lac* loop in vivo. *Nucleic Acids Res.* 38:8072–8082.
56. Czapla, L., J. Peters, ..., L. Maher. 2011. Understanding apparent DNA flexibility enhancement by HU and HMGB architectural proteins. *J. Mol. Biol.* 409:278–289.
57. Perez, P., N. Clauvelin, ..., W. Olson. 2014. What controls DNA looping? *Int. J. Mol. Sci.* 15:15090–15108.
58. Schneider, R., R. Lurz, ..., G. Muskheishvili. 2001. An architectural role of the *Escherichia coli* chromatin protein FIS in organising DNA. *Nucleic Acids Res.* 29:5107–5114.
59. Skoko, D., D. Yoo, ..., R. Johnson. 2006. Mechanism of chromosome compaction and looping by the *Escherichia coli* nucleoid protein Fis. *J. Mol. Biol.* 364:777–798.
60. Weber, S., A. Spakowitz, and J. Theriot. 2012. Nonthermal ATP-dependent fluctuations contribute to the in vivo motion of chromosomal loci. *Proc. Natl. Acad. Sci.* 109:7338–7343.
61. Chen, Y., J. Milstein, and J. Meiners. 2010. Protein-mediated DNA loop formation and breakdown in a fluctuating environment. *Phys. Rev. Lett.* 104:258103.
62. Swigon, D., B. Coleman, and W. Olson. 2006. Modeling the Lac repressor-operator assembly: the influence of DNA looping on Lac repressor conformation. *Proc. Natl. Acad. Sci. U S A.* 103:9879–9884.
63. Agrawal, N., R. Radhakrishnan, and P. Purohit. 2008. Geometry of mediating protein affects the probability of loop formation in DNA. *Biophysical J.* 94:3150–3158.
64. Czapla, L., M. Grosner, ..., W. Olson. 2013. Interplay of protein and DNA structure revealed in simulations of the *lac* operon. *PLoS One.* 8:e56548.
65. Corless, S., and N. Gilbert. 2017. Investigating DNA supercoiling in eukaryotic genomes. *Brief. Funct. Genomics.* 16:379–389.
66. Peters, J., V. Rao, ..., L. J. Maher. 2019. Dependence of DNA looping on *Escherichia coli* culture density. *Int. J. Biochem. Mol. Biol.* 10:32–41.
67. Champion, K., and N. P. Higgins. 2007. Growth rate toxicity phenotypes and homeostatic supercoil control differentiate *Escherichia coli* from *Salmonella enterica* Serovar Typhimurium. *J. Bacteriol.* 189:5839–5849.
68. Balke, V., and J. Gralla. 1987. Changes in the linking number of supercoiled DNA accompany growth transitions in *Escherichia coli*. *J. Bacteriol.* 169:4499–4506.
69. Mogil, L., N. Becker, and L. J. Maher. 2016. Supercoiling effects on short-range DNA looping in *E. coli*. *PLoS One.* 11:e0165306.
70. Fulcrand, G., S. Dages, ..., F. Leng. 2016. DNA supercoiling, a critical signal regulating the basal expression of the *lac* operon in *Escherichia coli*. *Sci. Rep.* 6:19243.
71. Higgins, N., and A. Vologodskii. 2015. Topological behavior of plasmid DNA. *Microbiol. Spectr.* 3:2.
72. Higgins, N. 2016. Species-specific supercoil dynamics of the bacterial nucleoid. *Biophysical Rev.* 8:113–121.
73. Kouzine, F., S. Sanford, ..., D. Levens. 2008. The functional response of upstream DNA to dynamic supercoiling in vivo. *Nat. Struct. Mol. Biol.* 15:146.
74. Peter, B., J. Arsuaga, ..., N. Cozzarelli. 2004. Genomic transcriptional response to loss of chromosomal supercoiling in *Escherichia coli*. *Genome Biol.* 5:R87.
75. Rui, S., and Y.-C. Tse-Dinh. 2003. Topoisomerase function during bacterial responses to environmental challenge. *Front. Biosci.* 8:256–263.
76. Liu, Y., V. Bondarenko, ..., V. Studitsky. 2001. DNA supercoiling allows enhancer action over a large distance. *Proc. Natl. Acad. Sci. U S A.* 98:14883–14888.
77. Yan, Y., F. Leng, and D. Dunlap. 2018. Protein-mediated looping of DNA under tension requires supercoiling. *Nucleic Acids Res.* 46:2370–2379.
78. Yan, Y., W. Xu, ..., L. Finzi. 2021. Negative DNA supercoiling makes protein-mediated looping deterministic and ergodic within the bacterial doubling time. *Nucleic Acids Res.* 49:11550–11559.
79. Podtelezhnikov, A., and A. Vologodskii. 1997. Simulations of polymer cyclization by Brownian dynamics. *Macromolecules.* 30:6668–6673.
80. Chirico, G., and J. Langowski. 1994. Kinetics of DNA supercoiling studied by Brownian dynamics simulation. *Biopolymers.* 34:415–433.
81. Chirico, G., and J. Langowski. 1996. Brownian dynamics simulations of supercoiled DNA with bent sequences. *Biophysical J.* 71:955–971.
82. Kalkbrenner, T., A. Arnold, and S. J. Tans. 2009. Internal dynamics of supercoiled DNA molecules. *Biophysical J.* 96:4951–4955.
83. Shusterman, R., T. Gavrinov, and O. Krichevsky. 2008. Internal dynamics of superhelical DNA. *Phys. Rev. Lett.* 100:098102.
84. Wedemann, G., C. Munkel, ..., J. Langowski. 1998. Kinetics of structural changes in superhelical DNA. *Phys. Rev. E.* 58:3537–3546.
85. Podtelezhnikov, A., and A. Vologodskii. 2000. Dynamics of small loops in DNA molecules. *Macromolecules.* 33:2767–2771.
86. Oberstrass, F., L. Fernandes, and Z. Bryant. 2012. Torque measurements reveal sequence-specific cooperative transitions in supercoiled DNA. *Proc. Natl. Acad. Sci. U S A.* 109:16.
87. Oberstrass, F., L. Fernandes, ..., Z. Bryant. 2013. Torque spectroscopy of DNA: base-pair stability, boundary effects, backbending, and breathing dynamics. *Phys. Rev. Lett.* 110:178103.
88. Mizuuchi, K., M. Mizuuchi, and M. Gellert. 1982. Cruciform structures in palindromic DNA are favored by DNA supercoiling. *J. Mol. Biol.* 156:229–243.
89. Beltran, B., D. Kannan, ..., A. Spakowitz. 2019. Geometrical heterogeneity dominates thermal fluctuations in facilitating chromatin contacts. *Phys. Rev. Lett.* 123:208103.
90. Fogg, J., A. Judge, ..., L. Zechiedrich. 2021. Supercoiling and looping promote DNA base accessibility and coordination among distant sites. *Nat. Commun.* 12:5683.

Biophysical Journal, Volume 121

Supplemental information

Coarse-grained modeling reveals the impact of supercoiling and loop length in DNA looping kinetics

Charles H. Starr, Zev Bryant, and Andrew J. Spakowitz

Supplementary Text

Coarse-Grained Modeling Reveals the Impact of Supercoiling and Loop Length in DNA Looping Kinetics

Charles H. Starr, Zev Bryant, and Andrew J. Spakowitz

1 Derivation of mean first passage time formula based on SSS theory

We treat loop formation as a diffusive process across the site-to-site distance reaction coordinate R under the influence of the free energy function $F(R)$. We choose a particular site-to-site distance R_T as the location of the transition state, such that states for which $R < R_T$ are "looped" and states for which $R_T < R < L$ are "unlooped". We assume that the system starts off unlooped, and becomes looped upon its first encounter with the transition state.

For such a system, the Green function $G^R(R', t|R, 0)$, which denotes the probability of transitioning from a starting site-to-site distance R at time $t = 0$ to a site-to-site distance R' at time t , is described by the following Fokker-Planck/Kolmogorov equation:

$$\left(\frac{\partial}{\partial t} - \Gamma^R \right) G^R(R', t|R, 0) = 0, \quad (1)$$

where

$$\Gamma^R = D \left(\frac{\partial^2}{\partial R^2} - \frac{\partial \beta F}{\partial R} \frac{\partial}{\partial R} \right). \quad (2)$$

In Eq. 2, D is the effective diffusion coefficient along the free energy profile $F(R)$.

Next, we define the survival probability $S(R, t)$ as the probability of remaining in any unlooped state $R' \in (R_T, L]$ at time t after starting from a particular unlooped state $R \in (R_T, L]$

$$S(R, t) = \int_{R_T}^L dR' (R')^2 G^R(R', t|R, 0). \quad (3)$$

By performing this same integration on Eq. 1, we see that the survival probability follows a similar dynamic equation as the Green function G^R , given by

$$\left(\frac{\partial}{\partial t} - \Gamma^R \right) S(R, t) = 0. \quad (4)$$

Taking the Laplace transform of this equation from t to the Laplace variable s leads to

$$s\tilde{S}(R, s) - S(R, t=0) - \Gamma^R \tilde{S}(R, s) = 0. \quad (5)$$

Because survival is guaranteed if no time has passed, $S(R, t=0) = 1$. Then, we set $s = 0$ and recognize that $\tilde{S}(R, s=0) = \int_0^\infty dt S(R, t)$ defines the average time $T(R)$ spent in the unlooped state before reaching the transition state R_T . Thus, we have a second-order differential equation for $T(R)$

$$\Gamma^R T(R) = D \left(\frac{\partial^2}{\partial R^2} - \frac{\partial \beta F}{\partial R} \frac{\partial}{\partial R} \right) T(R) = -1. \quad (6)$$

We define an absorbing boundary condition at the transition state [i.e. $T(R_T) = 0$] and a no-flux boundary condition at full extension [i.e. $\partial T(L)/\partial R = 0$].

Our task is now to solve this equation for $T(R)$ and to then compute the mean first-passage time as a Boltzmann-weighted average of $T(R)$ over the unlooped domain. By setting $\Phi = \frac{\partial T}{\partial R}$, we transform Eq. 6 into a first-order differential equation. We then multiply by an integrating factor $I = e^{-\beta F}$, leading to

$$e^{-\beta F} \left(\frac{\partial}{\partial R} - \frac{\partial \beta F}{\partial R} \right) \Phi = -\frac{e^{-\beta F}}{D}. \quad (7)$$

Next, we integrate Eq. 7, recognizing the left-hand side as

$$\frac{\partial}{\partial R} (e^{-\beta F} \Phi). \quad (8)$$

The no-flux boundary condition $\Phi(L) = 0$ lets us write

$$\Phi(R) = \frac{e^{\beta F(R)}}{D} \int_R^L e^{-\beta F(R'')} dR'' \quad (9)$$

We integrate once again to obtain $T(R)$. The absorbing boundary condition $T(R_T) = 0$ gives us

$$T(R) = \frac{1}{D} \int_{R_T}^R dR' \int_{R'}^L dR'' e^{\beta F(R') - \beta F(R'')} \quad (10)$$

Finally, we arrive at the mean first passage time from the unlooped state to the transition state by computing a Boltzmann-weighted average of $T(R)$ over all initial unlooped positions, such that

$$\begin{aligned} \langle T \rangle &= \frac{1}{DZ} \int_{R_T}^L dR \int_{R_T}^R dR' \\ &\quad \times \int_{R'}^L dR'' e^{-\beta F(R) + \beta F(R') - \beta F(R'')} \end{aligned} \quad (11)$$

where $Z = \int_{R_T}^L e^{-\beta F(R)} dR$ is the unlooped state partition function.

2 Supplementary Figures

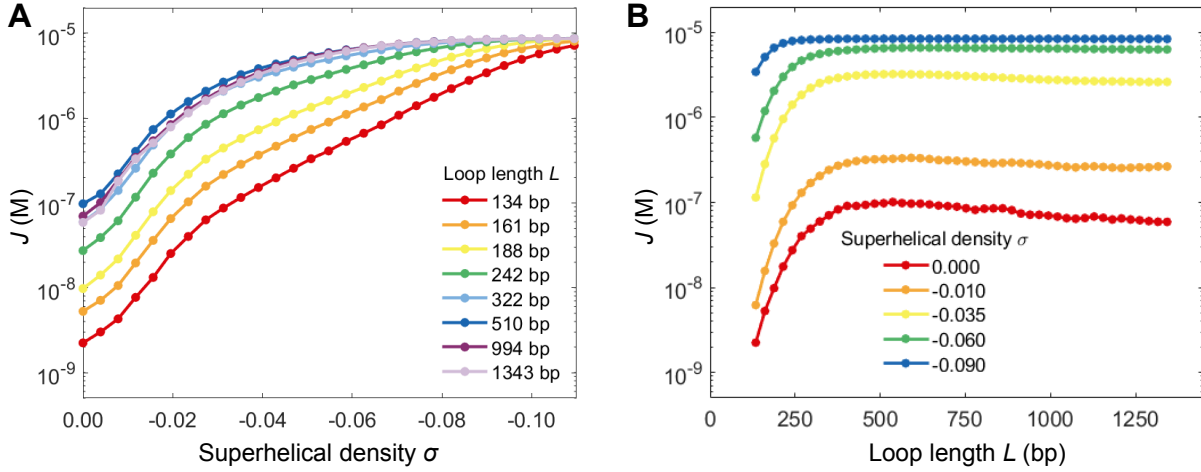


Figure S1: **J -factor versus superhelical density σ and loop length L .** These graphs are based on the same simulations ($L_{\text{circle}} = 2,686$ bp, 100 beads, $l_{\text{HC}} = 2$ nm) used to compute the $\langle T \rangle$ values of Figures 4A and 4C in the main text. **(A)** J versus σ curves for particular L values. **(B)** J versus L curves for particular σ values.

Overall, the response of the J -factor to supercoiling follows the same trend as the loop formation time $\langle T \rangle$. Supercoiling accelerates loop formation, mostly during the initial buckling transition with the response flattening out at higher superhelical densities. Shorter loops continue to accelerate with supercoiling even at higher superhelical densities. However, we observe subtle differences between J and $\langle T \rangle$ with respect to their supercoiling and loop length dependencies. For all loop lengths, the magnitude of the supercoiling-driven acceleration in looping is larger for J than for $\langle T \rangle$. In addition, J and $\langle T \rangle$ predict qualitatively different loop length dependencies at the highest $|\sigma|$ values we analyze. At $\sigma = -0.090$, for example, J predicts a minor acceleration in loop formation as loop length increases from the scale of one persistence length, while $\langle T \rangle$ predicts a minor deceleration (Figure 4C, blue curve).



Figure S2: **Randomly-chosen plasmid structures from our main simulation** ($L_{\text{circle}} = 2,686$ bp, 100 beads, $l_{\text{HC}} = 2$ nm) **at three superhelical densities**. Comparing structures from the center column to those in the right column, we see that plectonemic end loops become tighter and plectonemic stems become straighter as σ goes from -0.035 to -0.090 .

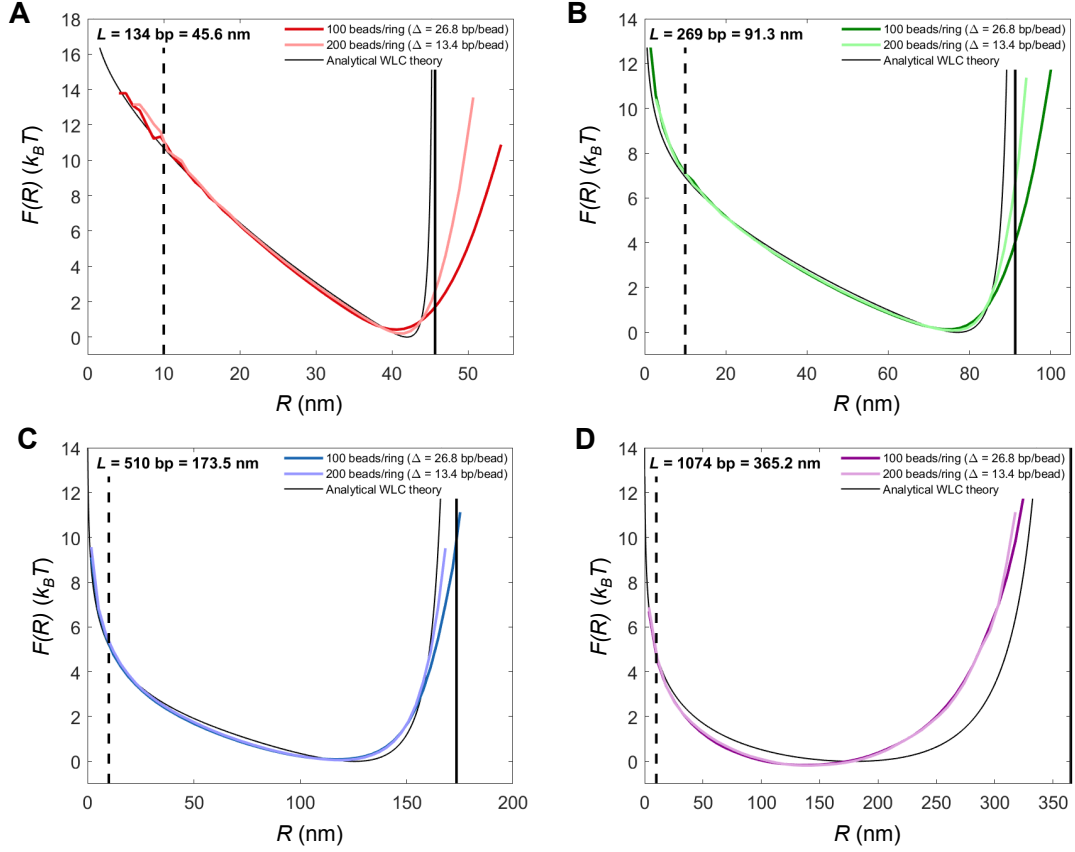


Figure S3: **Looping free energy profiles for $\sigma = 0.000$.** Colored curves represent $F(R)$ at $\sigma = 0.000$ from our main simulation ($L_{\text{circle}} = 2,686$ bp, 100 beads, $l_{\text{HC}} = 2$ nm) and an identical simulation with double the number of beads for four loop lengths (A,B,C,D: 5, 10, 19, and 40% of the total ring length, respectively). The black curve is the theoretical free energy of an ideal linear wormlike chain, computed using $F(R) = -k_B T \log(R^2 G(R))$, where $G(R)$ is the probability of the two chain ends being a distance R apart. We use an exact result for $G(R)$ derived in Ref. (5). The vertical dashed and solid lines denote the transition state location ($R = 10$ nm) and the fully-extended contour length ($R = L$), respectively.

Extensibility artifacts (states with $R > L$) are visible in Figures A-C, and it is clear that this artifact becomes less apparent for longer chain lengths and as the discretization length decreases. We do not expect a perfect match between simulations and theory because our simulations account for the additional energetic effects of self-repulsion and DNA's ring topology. However, at $\sigma = 0.000$, these effects are minor, especially for loops significantly shorter than the total ring length. Therefore, the close correspondence between simulation and theory across most of the R range in Figures A-C validates the accuracy of our simulations. In Figure D, the chain is 40% of the total ring length, and in this case, the simulation is biased towards shorter R relative to theory due to the topological constraint of ring closure.

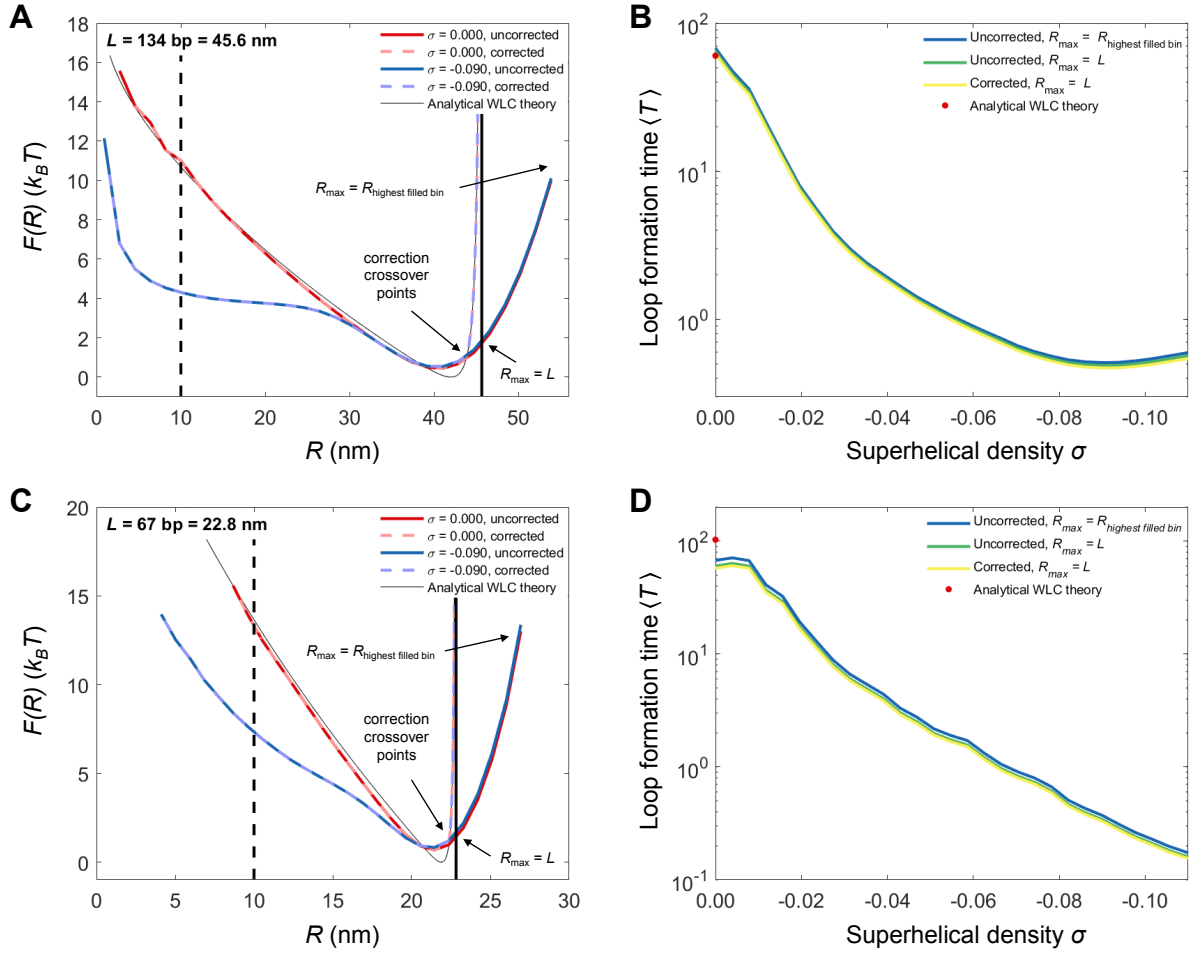


Figure S4: Extensibility artifact has a negligible effect on computed loop formation times. Many of our simulations contain over-extended loop conformations due to the lack of a finite extension constraint in the dssWLC model. As a result, we were forced to decide whether to use $R_{\max} = L$ as the reflecting boundary in Eq. 7 (main text), whether to use R of the rightmost sample-containing bin, or whether a correction would be necessary. To determine the numerical effect of these choices, we look at loops spanned by 5 bead-to-bead segments (a coarsely-discretized loop for which the extensibility artifact is extreme) from two simulations of the $L_{\text{circle}} = 2,686 \text{ bp}$ plasmid. Figures A and B display results from our main simulation (100 beads), in which 5 segments represent 134 bp, and Figures C and D display results from a 200-bead simulation, in which 5 segments represent 67 bp. **(A)** A rough correction procedure to enforce inextensibility for the 100-bead simulation. The dashed vertical line towards the left denotes the absorbing boundary at $R_T = 10 \text{ nm}$, and the solid vertical line towards the right denotes $R = L$. The solid lines show simulated and uncorrected free energy surfaces of the 134 bp loop. The distribution of simulated R values clearly extends into the $R > L$ range. To correct this behavior, we plot the theoretical free energy surface for a continuous chain of equivalent contour length, and we locate a "crossover point" as the point of intersection between the simulated and continuous free energy surfaces occurring to the right of the continuous free energy minimum. The corrected free energy surface then substitutes the continuous chain energies for the simulated energies for all R values to the right of the crossover point. **(B)** Loop formation time results for each of the choices depicted in A. The close match between curves demonstrates that free energies in the low- and intermediate- R range most strongly determine $\langle T \rangle$. We also show that $\langle T \rangle$ computed using the energy surface of the continuous linear chain closely matches our results for $\sigma = 0.000$. These energy surfaces differ dramatically near $R \sim L$, but are otherwise nearly identical. This confirms that the extensibility artifact has a minor numerical effect on kinetic results. **(C)** An identical correction procedure used to enforce inextensibility for the 200-bead simulation. **(D)** Loop formation time results for each of the choices depicted in C. Though we still observe a close match between curves, there is now a minor discrepancy between the $\sigma = 0.000$ data points and the linear WLC theory. To more accurately model loop lengths on this short scale, a more finely-discretized simulation would be required. However, such simulations require larger numbers of beads, reducing Monte Carlo sampling speed. Thus, in practice there is a lower limit to the length of a loop that can be accurately modeled using discretized simulations.

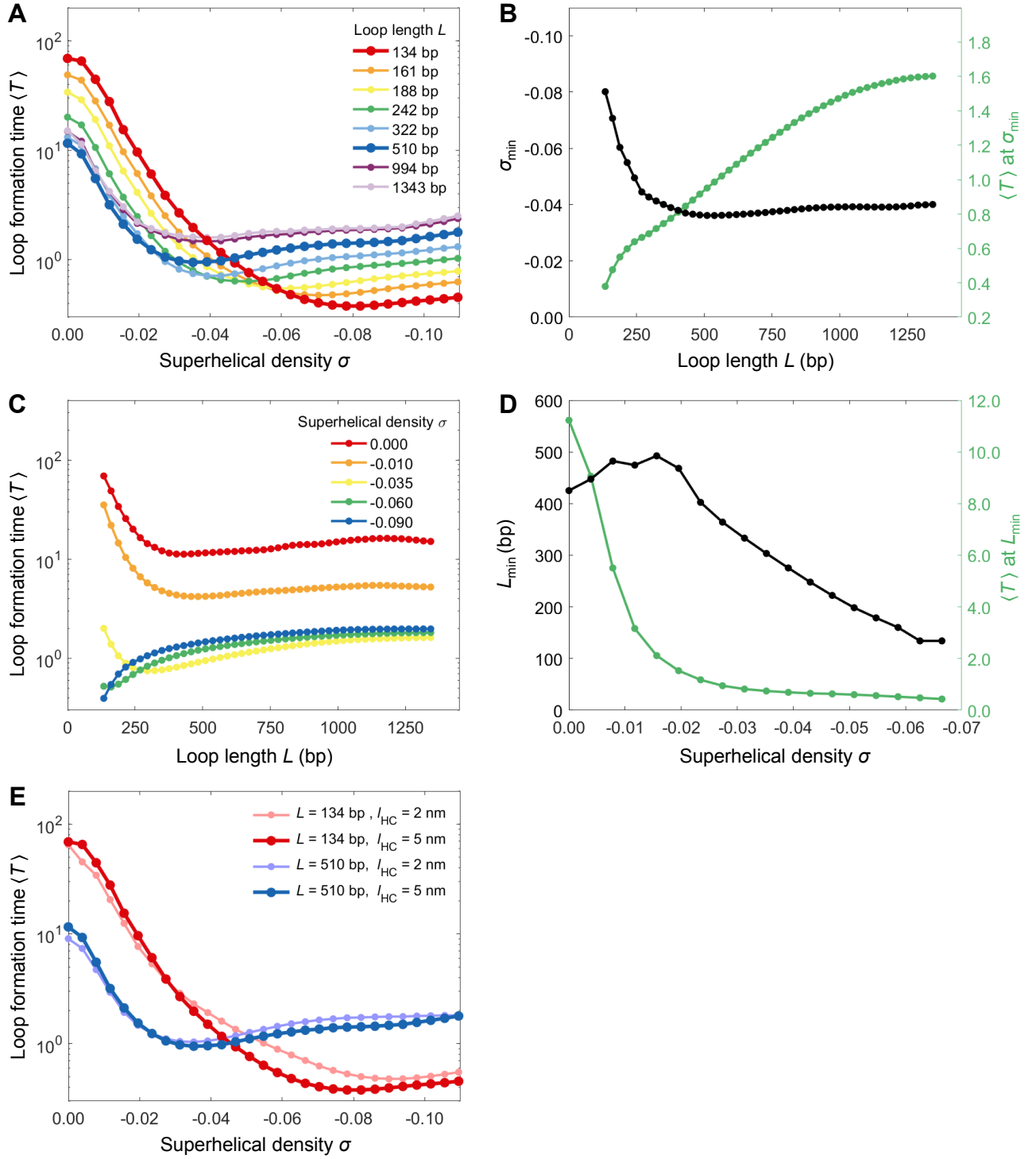


Figure S5: **Effect of the hard-core radius l_{HC} on loop formation times.** (A) $\langle T \rangle$ versus σ curves for particular L values from simulations with $L_{\text{circle}} = 2,686$ bp, 100 beads, and $l_{\text{HC}} = 5$ nm. The bold curves denote the two loop lengths featured in Figure E. (B) Location of the local minima in (A) as a function of L . (C) $\langle T \rangle$ versus L curves for particular σ values from the same simulations as in (A). (D) Location of the local minima in (C) as a function of σ . (E) Comparison of $\langle T \rangle$ versus L curves between $l_{\text{HC}} = 2$ nm and $l_{\text{HC}} = 5$ nm simulations for two loop lengths. In (B) and (D), minima are determined by fitting a parabola to the minimum data point and its two nearest neighbors.

Though the effect of l_{HC} is minor overall, it is most pronounced for the shorter loop at higher superhelical densities.

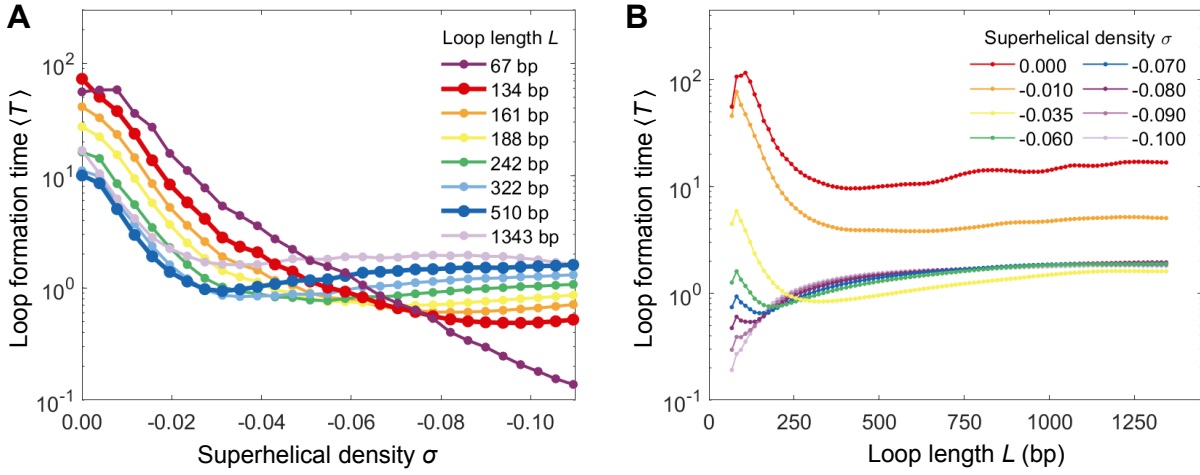


Figure S6: **Loop formation time $\langle T \rangle$ versus superhelical density σ and loop length L for a 2,686 bp ring simulated with 200 beads and $l_{\text{HC}} = 2$ nm.** (A) $\langle T \rangle$ versus σ curves for particular L values. (B) $\langle T \rangle$ versus L curves for particular σ values.

For $\langle T \rangle$ versus L curves showing local minima, we plot the locations of these minima as pale lines in Figure 4D of the main text. The additional spatial resolution gained in doubling the number of beads used to represent the ring allows us to resolve local minima at higher superhelical densities than in the 100-bead simulation of the main text. The leftmost data point in each curve ($L = 67$ bp) appears to deviate from the overall trend with respect to L , and therefore, we exclude these points when determining minima. This deviation likely results from an underestimate in $\langle T \rangle$ due to the extensibility artifact described in Figure S4C and S4D. This deviation may also reflect the fact that $\langle T \rangle$ must approach zero as $L \rightarrow R_T$. Local minima are determined by fitting a parabola to the minimum data point and its two nearest neighbors.

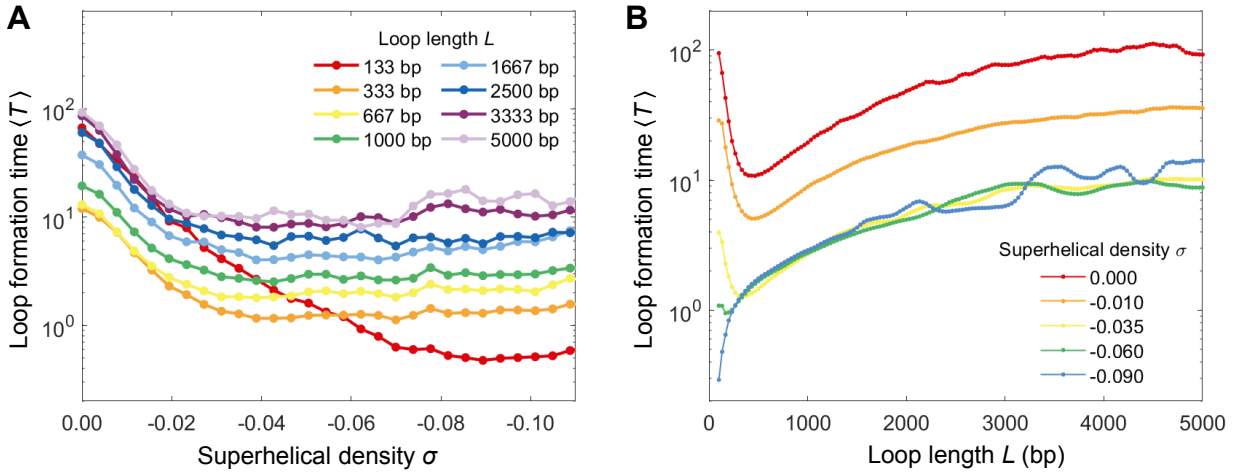


Figure S7: **Loop formation time $\langle T \rangle$ versus superhelical density σ and loop length L for a 10,000 bp ring simulated with 300 beads and $l_{\text{HC}} = 2$ nm.** (A) $\langle T \rangle$ versus σ curves for particular L values. (B) $\langle T \rangle$ versus L curves for particular σ values. For this larger plasmid, site-to-site distances can greatly exceed $R_T = 10$ nm, making a reduction in binwidth necessary to ensure the existence of at least one bin below R_T when constructing free energy surfaces. These kinetic results were constructed using a binwidth of $R/L = 0.01$.

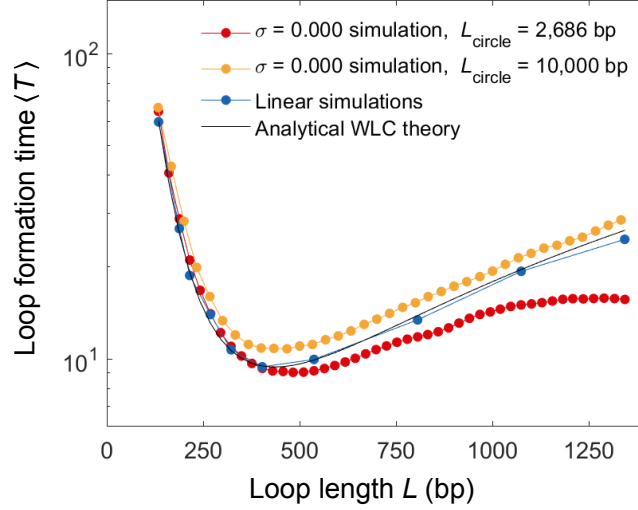


Figure S8: **Comparing kinetic results for simulated rings at $\sigma = 0.000$ to simulated and theoretical linear chains.** The red line depicts results from our main simulation ($L_{\text{circle}} = 2,686$ bp, 100 beads, $l_{\text{HC}} = 2$ nm). The orange line depicts results from a simulation with $L_{\text{circle}} = 10,000$ bp, 300 beads, and $l_{\text{HC}} = 2$ nm. Linear simulations are performed using the same parameters and procedure as our main ring simulation, except topological constraints and twist energetics are ignored. $\langle T \rangle$ predictions for the analytical WLC theory are computed by applying Eq. 7 from the main text to $F(R) = -k_B T \log(R^2 G(R))$, where $G(R)$ is the probability of the two chain ends being a distance R apart. We use an exact result for $G(R)$ derived in Ref. (5).

We display $\langle T \rangle$ versus L for loop lengths up to half the total length of the shorter ring (i.e. up to $L_{\text{circle}}/2 = 1,343$ bp). For $L < 500$ bp, $\langle T \rangle$ for the ring simulations are nearly equivalent to $\langle T \rangle$ for both simulated and theoretical linear chains, reflecting the minor effect of embedding short chains within rings. As L approaches half the total length of the shorter ring, the ring closure constraint biases energy surfaces towards smaller R (see Figure S3D), leading to slightly smaller $\langle T \rangle$ values. The loop lengths plotted here are small compared to half the total length of the longer ring (i.e. $L_{\text{circle}}/2 = 5,000$ bp), and therefore, $\langle T \rangle$ values for these simulations deviate less from our simulated and theoretical results for linear chains.

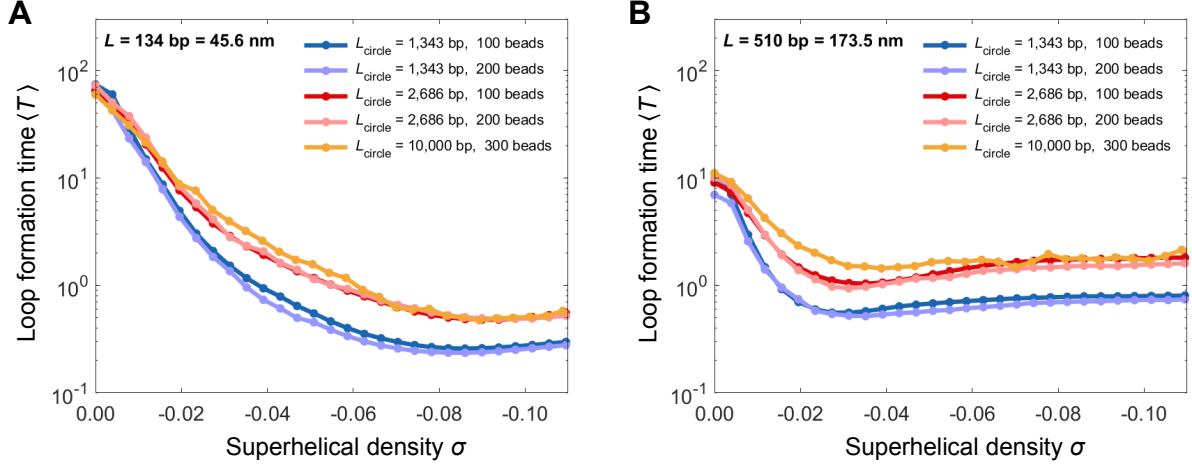


Figure S9: **Effect of ring length and discretization on loop formation times.** **(A)** $\langle T \rangle$ versus σ curves for $L = 134$ bp. **(B)** $\langle T \rangle$ versus σ curves for $L = 510$ bp. For the $L_{\text{circle}} = 10,000$ bp simulation, the chosen loop lengths do not correspond to an integer number of beads, so values are linearly-interpolated from the two nearest loop lengths available.

For each loop length L , we show curves for 3 ring lengths L_{circle} , with two levels of discretization shown for each of the two shorter rings ($L_{\text{circle}} = 1,343$ bp and $L_{\text{circle}} = 2,686$ bp). For the two shorter L_{circle} simulations, there is a negligible dependence of our results on the discretization length, justifying our use of a 100-bead simulation in the main text. Near $\sigma = 0.000$, all curves cluster near the same $\langle T \rangle$. However, at higher $|\sigma|$, our model predicts a slight dependence on L_{circle} . Aside from minor deviations (due in part to sampling error), $\langle T \rangle$ values for $L_{\text{circle}} = 10,000$ bp agree closely with those for $L_{\text{circle}} = 2,686$ bp. However, $\langle T \rangle$ values for $L_{\text{circle}} = 1,343$ bp tend to be roughly two-fold smaller than for either of the longer L_{circle} curves. This result differs from the L_{circle}^2 and L_{circle}^3 scaling of $\langle T \rangle$ predicted by Ref. (20) for branched and unbranched plectonemic structures, respectively. However, unlike the model in Ref. (20), our model captures the effect of supercoiling-driven shifts in the equilibrium distribution of structures on looping kinetics, without considering the explicit kinetic mechanism by which structures rearrange themselves. In theory, Brownian dynamics simulations capture both of these effects, though inefficient sampling relative to our Monte Carlo approach places practical limits on the span of loop lengths and superhelical densities that can be analyzed using these methods.

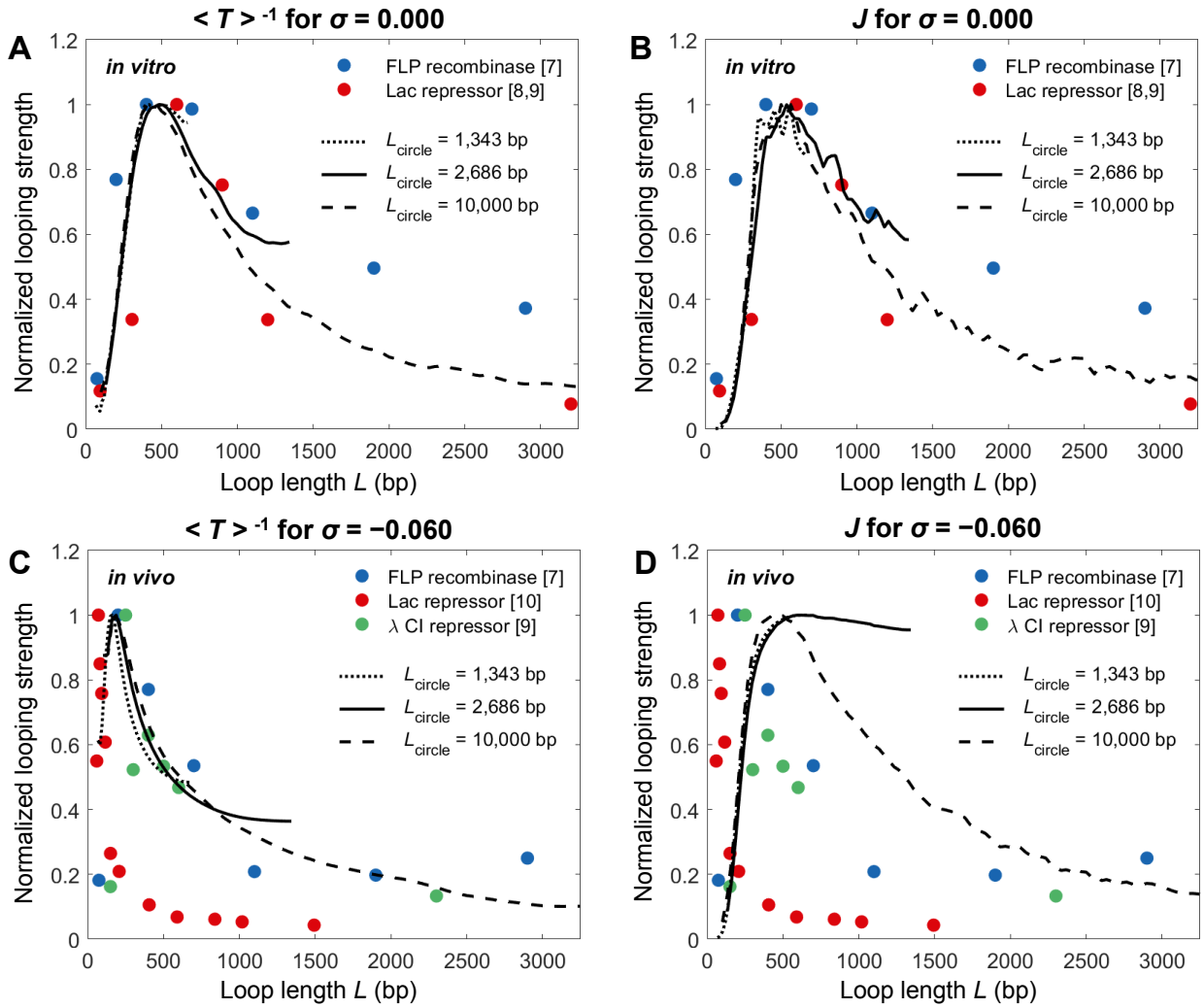


Figure S10: **Ring length has a small effect on the location of the peak loop length for $\langle T \rangle^{-1}$ and J .** In this figure, we display results from 100-bead simulations with $l_{HC} = 2$ nm for $L_{\text{circle}} = 1,343$ bp and $L_{\text{circle}} = 2,686$ bp. **(A,C)** $\langle T \rangle^{-1}$ versus L curves for three values of L_{circle} . In (A), we display results from $\sigma = 0.000$ simulations overlaid on experimental data points from in vitro looping assays. In (C), we display results from $\sigma = -0.060$ simulations overlaid on experimental data points from in vivo looping assays. **(B,D)** J versus L curves for three values of L_{circle} . In (B), we display results from $\sigma = 0.000$ simulations overlaid on experimental data points from in vitro looping assays. In (D), we display results from $\sigma = -0.060$ simulations overlaid on experimental data points from in vivo looping assays. As in Figure 5 of the main text, each series of data points or curve is normalized to its maximum to facilitate comparisons of looping peak locations.