

Single Molecule Investigations of Impacts from Mycobacterial Topoisomerase I C-terminal Deletions on Supercoil Relaxation

Dillon Balthrop ¹, Deepesh Sigdel ¹, Chunfeng Mao ¹, Yuk-Ching Tse-Dinh ^{2,3} and Maria Mills ^{1,*}

¹Department of Physics and Astronomy, [University of Missouri](#), Columbia, Missouri 65201, USA

²Department of Chemistry and Biochemistry, [Florida International University](#), Miami, Florida 33199, USA

³Biomolecular Sciences Institute, [Florida International University](#), 11200 SW 8th Street, Miami, Florida 33199, USA



(Received 29 October 2024; accepted 11 June 2026; published 6 July 2026)

Type IA topoisomerases relieve torsional stress in DNA by a strand-passage mechanism, using the strain in the DNA to drive relaxation. The topoisomerase IAs of the *Mycobacterium* genus have distinct C-terminal domains which are crucial for successful strand passage. We used single molecule magnetic tweezers to observe supercoil relaxation by wild type *Mycobacterium smegmatis* topoisomerase IA and two C-terminal truncation mutants. We recorded distinct behaviors from each truncation mutant. We calculated the free energy stored in the DNA as it is twisted under force to examine the differences between the proteins. Based on our results, we propose a modified model of the strand-passage cycle.

DOI: [10.1103/9tkx-w3zj](https://doi.org/10.1103/9tkx-w3zj)

I. INTRODUCTION

During various cellular processes torque is exerted on DNA causing torsional strain [1]. This strain can either twist the DNA about its helical axis or contort that axis in three dimensions, causing the DNA to wrap over itself, forming plectonemes. To describe this topology, the parameters twist (Tw) and writhe (Wr) are used. A DNA's Tw is the number of rotations it makes around its helical axis [Fig. 1(a)], whereas the number of three-dimensional crossings it makes with itself is called Wr [Fig. 1(b)]. Initially, Tw has lower energy than Wr . As Tw builds up Wr becomes more favorable, causing buckling. Tw and Wr can coexist and interconvert [2] [Fig. 1(c)], and their sum is called the DNA's linking number (Lk) (Fig. 1) [3]. When relaxed, DNA of arbitrary contour length L has a linking number Lk_0 . Each change in linking number (ΔLk) from Lk_0 is a supercoil. Rotating the DNA in the direction opposite to its natural right-handed helix decreases the linking number [4] creating negative supercoils. Rotating in the same direction increases Lk [4] creating positive supercoils.

The accumulation of excess supercoils renders DNA inaccessible, compromising cellular processes [3,5–7]. To manage DNA topology and relieve torsional stress, a mechanism exists in the form of enzymes called DNA topoisomerases [8–10]. Type IA topoisomerases relax supercoils via a strand-passage mechanism [11–13] [Fig. 2(a)] in which one strand of the DNA duplex is cleaved (gate strand or G strand) and the

opposing strand is passed through the break (transport strand or T strand). The broken G strand is held open by a protein-mediated DNA gate [11]. This process removes supercoils in steps of 1 ($\Delta Lk = 1$) and does not require an input of energy. Instead, the process is driven by thermal energy and torsional strain in the DNA molecule. Because these enzymes require a single stranded region to function, they typically relax negative supercoils which are more prone to base pair separation, also called melting [2,14]. Type IA topoisomerases are ubiquitous but there exists structural diversity between organisms. The N-terminal catalytic core (domains 1–4 or D1–D4) remains highly conserved from the archetypal *Escherichia coli* topoisomerase IA and binds to the G strand [15–17]. This core forms a “padlock” shape and contains the active site of the protein, which is responsible for reversibly cleaving the G strand [Fig. 2(b)]. Most type IA topoisomerases have additional C-terminal domains (CTDs), which vary in structure and may bind to the T strand [18,19]. The topoisomerases of mycobacteria, such as the disease-causing *Mycobacterium tuberculosis*, have four CTDs [Fig. 2(b)] which are essential for strand passage [18–20], though their precise dynamics are unclear. The mycobacterial topoisomerase IA CTDs are not present in human topoisomerases, making those domains potential targets for antibiotics [21,22].

Previous work has revealed the general role of the of the domains and their impact on supercoil relaxation [18,19,23]. *Mycobacterium smegmatis* TopI (MsmTopI), the focus of this study, is a common model for *Mycobacterium tuberculosis* TopI. MsmTopI has four C-terminal domains (D5–D8) and a 27-residue-long, basic, and unstructured tail [Fig. 2(b)]. These CTDs have high affinity for ssDNA binding [18,19]. Bulk biochemical experiments have shown that sequential deletions of MsmTopI CTDs resulted in exponentially diminishing relaxation activity [19]. Others have also shown that removing basic portions of the CTDs reduces activity [18]. The same truncated mutants used by Cao and colleagues [19]

*Contact author: k.maria.mills@missouri.edu

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](#) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

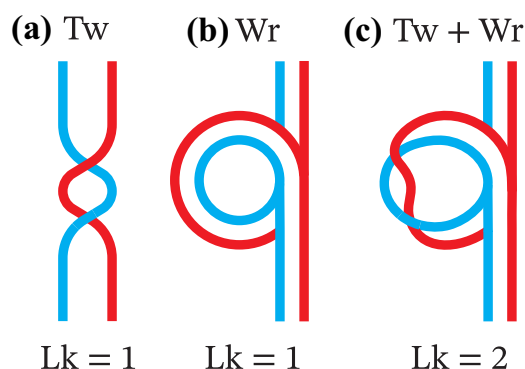


FIG. 1. Visual explanation of twist (Tw) and writhe (Wr). Each colored line is one strand of ssDNA. Two strands together form dsDNA. (a) One 360° rotation about the long axis is a twist. (b) One three-dimensional self-crossing of the full DNA backbone is a writhe. (c) Twist and writhe may exist simultaneously.

are investigated in this work. The first truncated mutant, 909t MsmTopI (909t), lacks the 27-residue, basic, unstructured tail, and terminates at the 909th residue at the end of D8. The second truncated mutant, 839t MsmTopI (839t), lacks both the tail and D8. Our work aims to closely examine how these deletions affect the relaxation process on a single molecule scale.

Single molecule techniques are ideal for probing the details of such protein mechanisms [24–28]. Single molecule magnetic tweezers (MT) are commonly used to analyze protein-nucleic acid interactions due to their ability to record and observe data in real time, capture multiple signals at once, and the small amount of perturbation being applied directly to the system [29–32]. MTs consist of an inverted microscope with a pair of overhanging permanent magnets. The magnets can be rotated clockwise (+) or anticlockwise (–) to introduce supercoils (SCs) to the DNA. The magnets may also move up or down to vary the force applied to the DNA (Fig. S1 in the Supplemental Material [33]). Overall, we observed distinct differences between all three enzymes. Ultimately, to link the loss in activity seen in bulk experiments and individual characteristics seen in single molecule experiments, we looked to the DNA's changing energy as supercoils were relaxed and what possible steps in strand passage could contribute to the observed behavior.

II. METHODS

A. Magnetic tweezers measurements

Measurements were performed on a combined magnetic tweezers–TIRF instrument purchased from MadCityLabs. Sample cells used for experiments were prepared as by Mao and Mills [34]. Experiment buffer solution was 50 mM Tris HCl, 100 mM K-Glut, 1 mM Mg-Glut, 0.01% BSA. An average force calibration versus magnet position for $1\ \mu\text{m}$ MyOne magnetic beads (Invitrogen) was obtained using Brownian motion analysis of 6 kb DNA tethers. The coilable DNA molecules were prepared as described by Seol and Neuman; the 11 kb DNA details are provided parenthetically [28]. In experiments with 6 kb DNA (11 kb DNA), a force range of

0.4–0.5 pN (0.6–0.7 pN) was applied to the bead with the magnetic tweezers to keep the DNA extended and away from the surface and causing slight base pair separation. Fields of view with multiple usable tethers were first identified. Tethers were considered usable if DNA could be supercoiled by rotating the overhead magnets and fell within a length range of $1.2\text{--}1.6\ \mu\text{m}$ ($2.8\text{--}3.2\ \mu\text{m}$) when relaxed. This length range was determined by considering the possible number of base pairs bound to the magnetic bead and cover slip. In other words, for a 6 kb tether, there is a 500 bp range which may bind to the magnetic bead, and another 500 bp which may bind to the flow cell. Variation of the attachment geometry within these ranges can decrease the tether's effective length from the ideal 6 kb. The extension change after the DNA buckles is approximately 50 nm/turn. The magnets were then completely raised to avoid adding stress to the tethers while the desired concentration of enzyme was added. Enzymes were purified as described by Cao *et al.* [19]. Once equilibrated, the force was set back to the prior value. The z position of the bead was tracked over time to monitor the changes in extension of the DNA tether. If the tether was relaxed to its original height, more supercoils would be introduced. This relaxation and rewinding would continue until the DNA became unusable, i.e., would not recoil, were nicked, or detached from the surface (Fig. 3). All data were collected at 60 Hz at room temperature ($21^\circ\text{C} \pm 2^\circ\text{C}$). The image tracking was accomplished with a modified labview program based on a program provided by Professor Michael Poirier of Ohio State University.

B. Data analysis

Experimental data were first loaded into IGOR PRO 9 and subjected to preprocessing, where it was decimated by a factor of 2 to reduce analysis times. It was then exported and analyzed using a custom PYTHON program which uses Student's t -test to identify run starts, and stops, in data traces. The time between steps is also recorded as the dwell time. The base of this program was provided by Dr. Ian Morgan (NIH, NHLBI). It then performs linear fits to the detected regions, and we pulled the relevant quantities from the fits. The algorithm implements multipass thresholding to avoid detection of events which are below our resolution limits. Once values were obtained, they were either further analyzed using PYTHON or binned in histograms using IGOR PRO 9. If multiple distributions were seen, the data would be put through another custom PYTHON algorithm which implements information criterion and K -means clustering to determine the optimal number of Gaussian distributions that fit the data.

C. DNA constant determinations

The method described here is based on calculations using a model developed by Marko [2], and expanded on in [35]. Relaxed DNA has a bend persistence length (A) which is the minimum length of DNA required for it to bend, and a twist persistence length (C) which is the length of DNA typically spanned by 1 Tw. C is also referred to as twist stiffness. The twist persistence length of plectonemic DNA is different than C and is denoted as P . Similarly, the twist rigidity of stretched

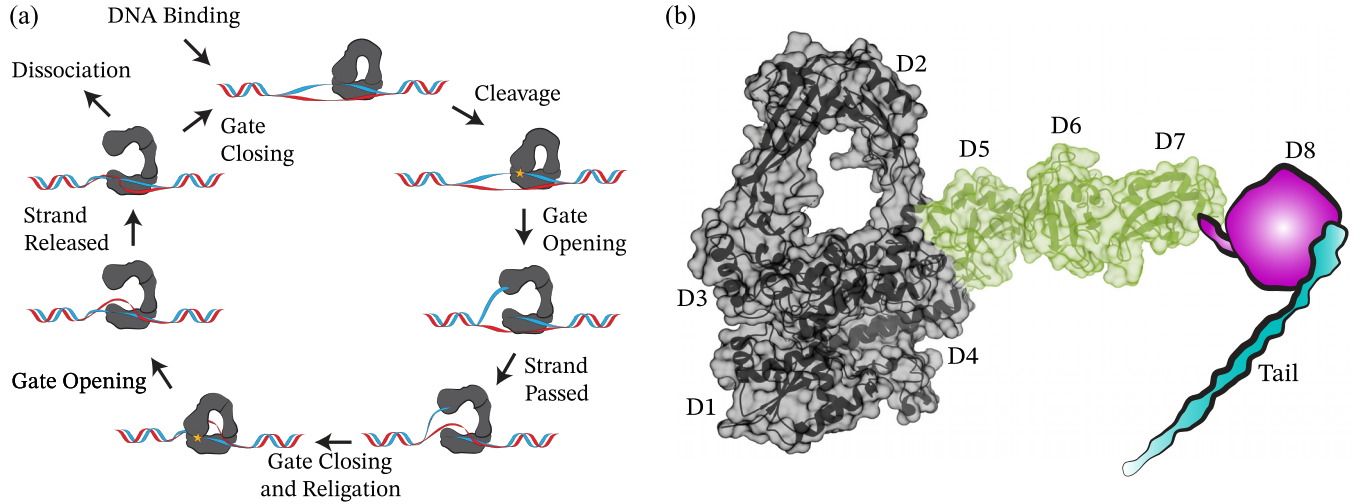


FIG. 2. (a) Current proposed strand-passage mechanism for type IA topoisomerases. The C-terminal domains are not shown. (b) Crystal structure of 839t MsmTopI deletant (PDB: 6PCM [19]) with missing domains sketched in. The N-terminal catalytic core is represented in gray. The domains are colored by presence in each mutant. Domains in the 839t mutant are D5–D7. The domains present in the 909t mutant are D5–D8. Throughout this work, data and schematic coloration are the same.

DNA under a force f is defined to be c_s , which governs the amount of twist that can be absorbed into writhe and follows the expression

$$c_s = c \left[1 - \frac{C}{4A} \left(\frac{k_B T}{A f} \right)^{1/2} \right]. \quad (1)$$

The presence of some constants makes it convenient to rewrite C and P as $c = k_B T C \omega_0^2$ and $p = k_B T P \omega_0^2$, respectively ($k_B T = 4.1$ pN nm). The supercoiling density,

$\sigma \equiv \Delta Lk / Lk_0$, is used in place of the linking number for calculation since it is a normalized quantity.

There are three stages of DNA supercoiling, each with independent expressions for their free energy. First, all ΔLk is added in the form of Tw. The free energy per length of this region is $\mathcal{S}(\sigma; f)$. Once the supercoiling density reaches a critical value, σ_s , the DNA buckles, entering the second stage and introducing the coexistence of Tw and Wr. This region's free energy per length is denoted as $\mathcal{F}(\sigma; f)$. The addition of Wr can be thought of as an exchange of extension for structural stability which keeps the torque constant. As σ continues to increase, it approaches another critical value, σ_p , and enters the third stage, where all Tw is converted to Wr, and the DNA is all plectonemic. In the case of stages 2 and 3 where the supercoils are negative, considerable DNA melting can occur at forces greater than 1 pN. Because our experiments were done at $f \simeq 0.45$ pN for 6 kb DNA and $f \simeq 0.65$ pN for 11 kb DNA, our negatively supercoiled DNA is melted enough to have single stranded regions for the enzymes to bind, but not experience large extension changes from denaturation. Thus we used the equations for positive supercoiling which are simpler. Simplifying further, we only needed to consider the first and second stages, because once the DNA reaches σ_p it would be coiled all the way to the surface, and there would be no measurable extension change. The equations describing free energy per length of the two regions of interest, respectively, are

$$\mathcal{S} = -g + \frac{c_s}{2} \sigma^2, \quad (2)$$

$$\mathcal{F} = -\frac{g}{1 - \frac{p}{c_s}} + \left[\frac{2pg}{(1 - p/c_s)} \right]^{1/2} |\sigma|. \quad (3)$$

The quantity g represents the free energy per length of stretched, freely swiveling DNA and follows the expression

$$g = f - \left(\frac{k_B T f}{A} \right)^{1/2}. \quad (4)$$

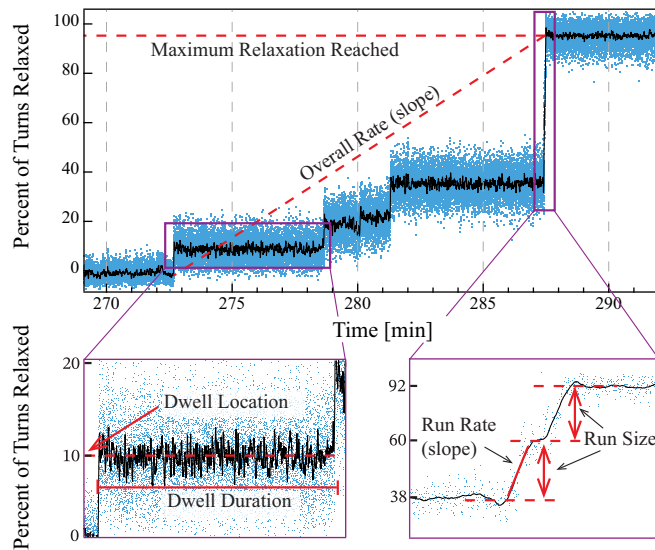


FIG. 3. Representative signal trace of negative supercoil relaxation by wild type *Mycobacterium smegmatis* topoisomerase IA, depicting a full relaxation event consisting of several smaller runs. The quantities of interest are labeled and two zoom-ins show properties of individual runs which were analyzed. The trace is part of a series of consecutive relaxation events which occurred during an experiment.

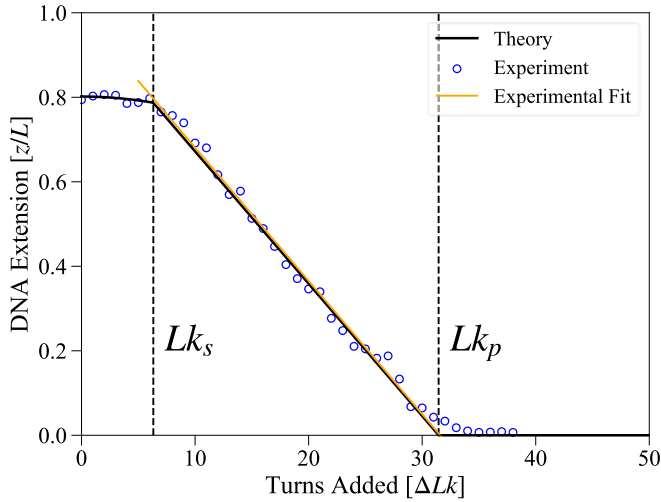


FIG. 4. Representative hat curve taken from experiment with a linear fit performed on the range between Lk_s and Lk_p . The y axis is the DNA extension per length, a normalized quantity where L is the DNA contour length. Experimental hat curves are used for comparison with the theoretical model employed in the main text to identify proper parameters for the model. All hat curves are taken in the absence of protein.

Similarly, c_s , c , and p all have units of free energy per length. Accurate values of A , C , and P depend on experimental conditions like salt types and concentrations. We determined the parameters for our system by taking several DNA extension curves, or hat curves (Fig. 4). Hat curves demonstrate the DNA extension change as supercoils are added by rotating the overhead magnets. Because the derived equations allow for the computation of theoretical extension curves, we optimized each parameter in computation until we saw agreement between the theoretical prediction and our experiments. The predicted curves (Fig. 4) were calculated by taking the negative of the partial derivative with respect to force, $-\partial/\partial f$, of Eqs. (2) and (3). The derivatives are representative of the DNA's extension per length, a normalized quantity, denoted as z/L where L is the DNA contour length, and are valid over the previous ranges of σ (Fig. 4). The results are

$$-\frac{\partial \mathcal{S}}{\partial f} = \frac{g}{2f} + \frac{1}{2} - \frac{c - c_s}{4f} \sigma^2, \quad (5)$$

$$-\frac{\partial \mathcal{F}}{\partial f} = \left[\frac{1}{2f\xi} - \frac{\sigma}{4f} \sqrt{\frac{2p}{g\xi}} \right] (f + g) + \frac{Ccp}{8Ac_s^2 f^2 \xi^2} \left[\sigma \sqrt{\frac{pg\xi}{2}} - g \right] (f - g), \quad (6)$$

where $\xi \equiv 1 - p/c_s$.

Once the parameter values are known, one can simply follow the formulation to obtain the equations of free energy which govern the three regimes of supercoiling density. As stated previously, we may disregard the third regime because, experimentally, the DNA would be fully supercoiled to the surface at that point, rendering no extension change.

III. RESULTS

A. Supercoil relaxation

We measured the DNA relaxation activity of MsmTopI by an established MT assay [11,29,36,37]. Functionalized 6 kb DNA tethers were attached on one end to the bottom cover slip of a flow cell and on the other end attached to a paramagnetic bead. The tethers were held at a constant force of 0.45 pN. This force causes slight melting of negatively supercoiled DNA and creates a single stranded binding region for the enzyme. The tethers were negatively supercoiled ($-SC$) by 30 turns in the presence of 1 nM WT MsmTopI. The z positions of the beads were recorded in real time as the enzyme relaxed the supercoils. Once the tethers were relaxed, they would then be recoiled, and the process would be repeated. Over the course of a single relaxation event there were typically multiple processive runs interspersed by pauses, or dwells, of varying durations (Fig. 3). The data were then fed to a custom PYTHON algorithm which implemented Student's t-test to identify the start and end points of each relaxation run and parameters of interest were calculated. For consistency, a threshold of 0.33 s between relaxation runs needed to be met for a dwell to be considered. It is important to note that we cannot claim to know that the same protein is responsible for all relaxation steps in the same event. However, we conducted the experiments at 1 nM because it was the minimum working concentration to observe activity. This maximizes the chances that there is a single enzyme per relaxation step. The results were plotted as histograms which were then fit with the appropriate distributions to obtain statistical quantities (Fig. 5). Experiments and analysis were performed identically for the mutants. For each enzyme variant, there was a range of 5–30 events per experiment, with a minimum of 10 experiments performed. In order to provide more granularity at higher force during supercoiling, we also performed experiments using 11 kb tethers at 0.65 pN. The only experimental difference between the groups is that the 11 kb DNA requires approximately 50–60 turns to be fully supercoiled.

While type IA topoisomerases are known to change the DNA linking number by discrete steps of one per kinetic cycle [38], they are also known to stochastically complete more cycles in a processive burst or run [32,36]. In our observations (Table I) the wild type (WT) enzyme relaxed 4.7 ΔLk , on average, per run with an average run rate of 7.9 $\Delta Lk/s$ [Figs. 5(a) and 5(b)]. The 909t mutant exhibits a notable preference for longer average run sizes of 8.7 ΔLk , but a slower run rate at 3.3 $\Delta Lk/s$. Run sizes for the 909t also spanned a wide range, with the enzyme relaxing up to 25 turns in a single run in some instances. The 839t has a short average run of 3.2 ΔLk , but works at a much faster run rate of 12.2 $\Delta Lk/s$. Considerable variance was observed in dwell times between relaxation events [Fig. 5(c)]. The WT dwell times averaged 6.59 s between runs with frequent instances of dwells exceeding 1 min. Analysis of dwell positions for the WT revealed a preference for dwelling primarily within the first and last 20% of relaxed turns [Fig. 6(a)]. However, there was no correlation between where a dwell occurred and its duration. In fact, the WT maintained relatively consistent pacing regardless of supercoil density. Average dwell times for the 909t mutant were longer at 8.14 s and more likely to

TABLE I. Overview of experimental results. Values are from fits to histograms, and errors are from 2 standard deviations. Counts of each variable are reported as n_{rs} , n_{rr} , and n_{dw} , respectively. N is the count for the total number of events.

Protein	Run size (ΔLk)	Run rate ($\Delta Lk/s$)	Dwell time (s)	Overall rate ($\Delta Lk/s$)
Wild type	4.73 ± 0.70 , $n_{rs} = 549$	7.9 ± 0.1 , $n_{rr} = 529$	6.59 ± 1.10 , $n_{dw} = 552$	0.34 ± 0.08 , $N = 105$
909t	8.74 ± 1.60 , $n_{rs} = 421$	3.3 ± 0.1 , $n_{rr} = 411$	8.14 ± 3.00 , $n_{dw} = 407$	0.67 ± 0.25 , $N = 159$
839t	3.21 ± 0.80 , $n_{rs} = 1217$	12.2 ± 0.1 , $n_{rr} = 1217$	6.14 ± 0.90 , $n_{dw} = 1217$	0.52 ± 0.14 , $N = 128$

occur with longer durations as more of the supercoils were removed. The 839t displayed dwell durations shorter than the WT with an average value of 6.14 s but has more dwells overall due to its smaller run sizes. Interestingly, the 839t does show some instances of prolonged dwells near Lk_0 , but the large number of dwells overall renders them statistically insignificant.

Increases of dwell times with dwell positions can be contextualized by the overall relaxation rate, which is defined to be the change in Lk per second over the time of a full relaxation event. The WT's overall rate was determined to be approximately $0.34 \Delta Lk/s$ [Table I, Fig. 5(d)]. The 909t overall relaxation rate is $0.67 \Delta Lk/s$. This is despite its exceedingly long dwell times at $Lk \approx Lk_0$. The 839t mutant achieves an overall rate of $0.52 \Delta Lk/s$, surpassing the WT but remaining under the 909t mutant's rate. Qualitatively speak-

ing, the 909t relaxes more turns without stopping, reducing the impact of its lower run rate by eliminating extra dwells, and slows down dramatically as $Lk \rightarrow Lk_0$. In contrast, the 839t relaxes in rapid and short bursts, but is plagued by many dwells of varying duration which pull its overall rate below the 909t's.

For the experiments which used 11 kb tethers, we applied a higher force of 0.65 pN. This higher force affects the amount of strain being put on the DNA, which will in turn affect the topoisomerase activity. The calculable energies support this idea, showing that the available energy in the DNA is theoretically higher throughout the length of the 11 kb tether (Fig. S3 in the Supplemental Material [33]). In fact, we observed what appears to be a convergence to some common set of relaxation results. The run sizes for the WT, 909t, and 839t were extremely similar at $4.01 \pm 0.91 \Delta Lk$, $4.44 \pm 0.75 \Delta Lk$, and $3.9 \pm 0.56 \Delta Lk$, respectively. The general trend seen before is still there but they are all now within error of one another. The run rates also became very similar with the original trend still being the same. The WT, 909t, and 839t had average run rates of $6.5 \Delta Lk/s$, $6.2 \Delta Lk/s$, and $6.7 \Delta Lk/s$, respectively, at 0.65 pN. Interestingly, the previous trends are shaken by the dwell times. On the whole, they became shorter, most dramatically being the 909t, with its dwell times decreasing to 1.10 ± 0.12 s. The WT and 839t also saw decreases to 3.62 ± 0.65 s and 2.86 ± 0.86 s. Interestingly, even though the dwells were generally shorter, there was still a tendency for the 909t and 839t to dwell longer as more turns were removed from the DNA [Fig. 6(b)]. The impact on average dwell times carries an explicit impact onto the resulting overall rates for each enzyme variant. However, it appears that only the magnitudes of their differences were changed, because we observed that the overall rates also carried the same trend seen at 0.45 pN. The WT, 909t, and 839t had overall rates which were generally slower, averaging at $0.14 \Delta Lk/s$, $0.48 \Delta Lk/s$, and $0.18 \Delta Lk/s$. Moreover, the longer dwells which occur for the 909t and 839t as more turns are relaxed pull down the averages. It is worth noting the sample sizes for the 11 kb are lower than that for the 6 kb, ranging from 20 to 70 total events per protein. Generally speaking, it seems that increasing the force causes the increased strain in the DNA to somewhat equalize the activity of the enzymes.

Given that the 909t was expected to be $\sim 1/2\times$ as effective as the WT, and the 839t $\sim 1/18\text{--}36\times$ based on previous results from bulk assay of relaxation of supercoiled plasmid [22], our results were initially puzzling. However, the overall rate only encapsulates the rate from start to stop of activity, and not all events ended with the DNA completely relaxed. We analyzed each protein's stopping Lk [Fig. 7(a)], defined as the ΔLk from fully relaxed after 30 min of no activity on the tether. Unimodal Gaussian distributions describe both the WT

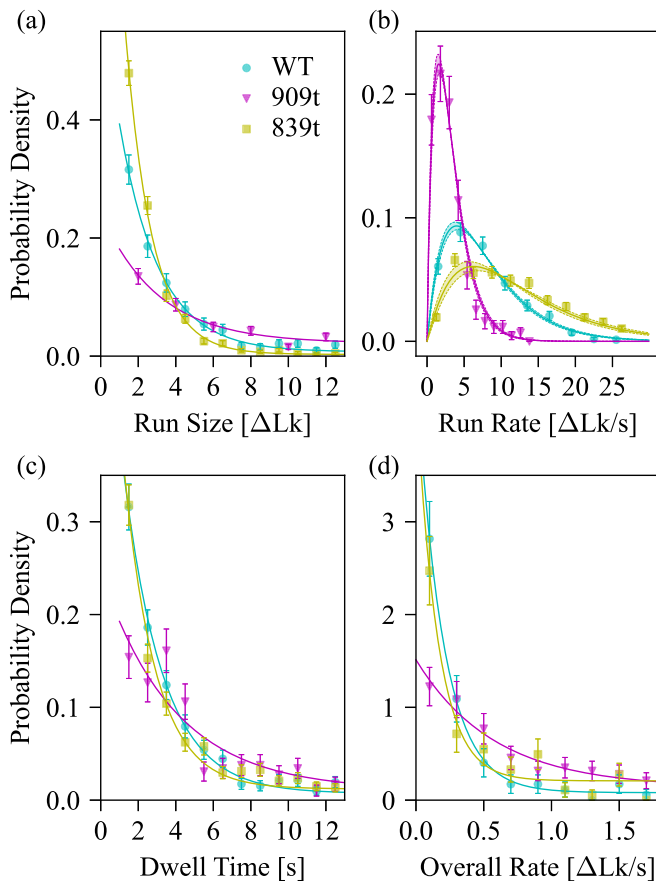


FIG. 5. Histograms showing the distributions for different properties that were analyzed. All data were normalized to a probability density. (a) Histogram of run size. (b) Histogram of run rate. (c) Histogram of dwell times. (d) Histogram of overall rate.

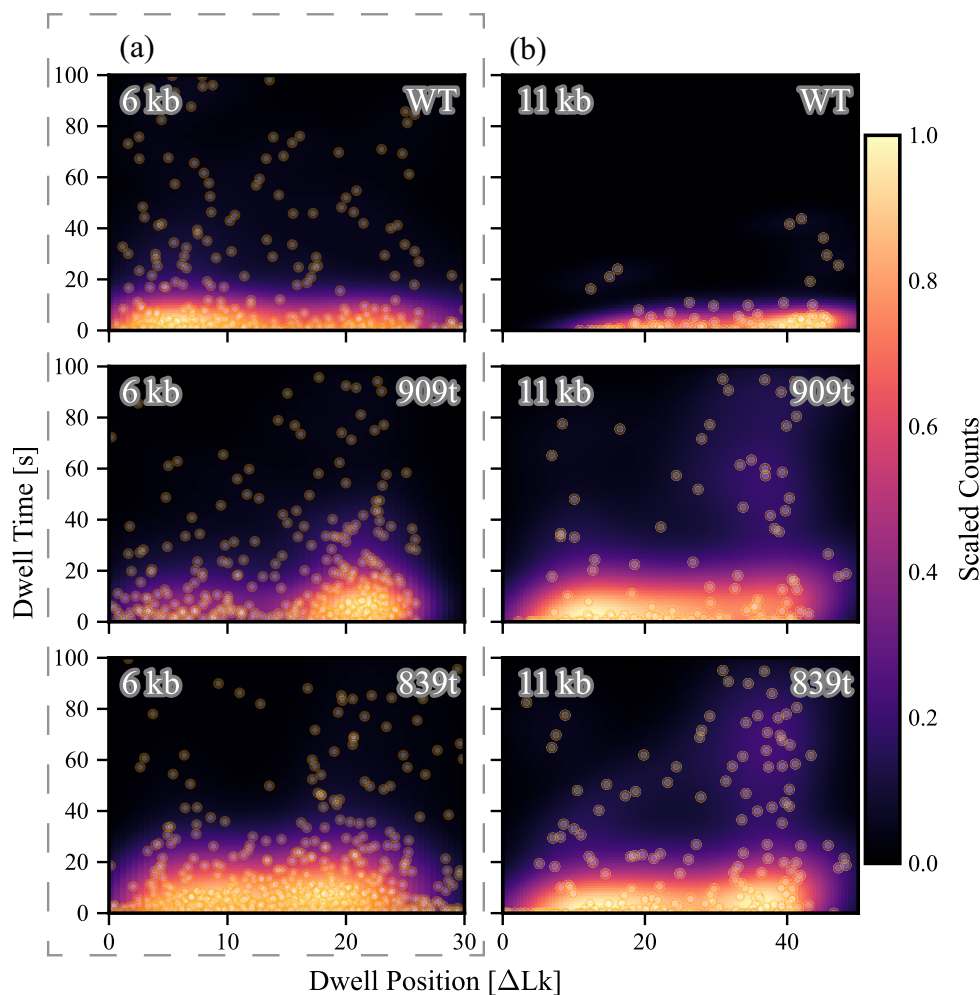


FIG. 6. Correlation heat maps for dwell times and dwell positions of each enzyme on (a) 6 kb tethers at 0.45 pN and (b) 11 kb tethers at 0.65 pN. Points which are high on the y axis are longer dwells. Points which occur further on the x axis represent dwelling after more supercoils were resolved from the substrate. Circles mark the individual points which make up the heat map for visual clarity in low concentration areas. Counts for the 6 kb tethers are found in Table I. Counts for the 11 kb tethers are 102, 179, and 343 (WT, 909t, 839t).

and 909t, centered at 8 and 10 ΔLk , respectively [Fig. 7(a)]. The WT's distribution was broad, ranging from 0 to 16 ΔLk remaining, but the 909t's was tighter, only covering from 5 to 15 ΔLk remaining. Surprisingly, the 839t displayed a bimodal Gaussian distribution with peaks centered at 11 and 4 ΔLk remaining, but cumulatively, it spans the same range as the WT.

The differences in stopping positions illustrate that the loss of the CTDs affects strand passage differently depending on how much supercoiling has been removed. In contrast to the WT, the 909t struggles to relax turns near Lk_0 . This suggests that the 27-residue tail plays an important role in supercoil removal as Lk approaches Lk_0 . This could be due to decreased binding to the T strand and lower energy remaining in the DNA for relaxation. Lysine and arginine residues present in the tail imply strong electrostatic interaction with the T strand and give credibility to these possibilities. Given that the proteins use the torsional strain in the DNA to perform relaxation, it is likely that the DNA buckling transition plays a role in all these distributions. Further, this energy deficit could be exacerbated in the 839t that has one fewer C-terminal domain

for binding to the T strand. To provide more information on the supercoil relaxation at the buckling transition, we turned to the 11 kb DNA tethers. The longer DNA requires more turns to overcome the buckling transition, which allowed us to use a higher force and verify the impact on relaxation at values near Lk_0 . The results for the 11 kb DNA echo those of the 6 kb. We saw continued diminished relaxation across the variants corresponding to the removal of enzyme domains. To probe the connection in deeper detail, we looked to the established model by Marko [2] which connects the DNA free energy to its level of supercoiling under force.

B. Energetic connection

For a brief overview of the theory used in this section, see Sec. II. In summary, we calculated the free energy per length of the DNA. By measuring extension changes when supercoiling the DNA in the absence of protein, one can compare the obtained curve to a theoretical prediction (Fig. 4). This comparison allows for the extraction of certain mechanical properties of the DNA which are then used to compute the

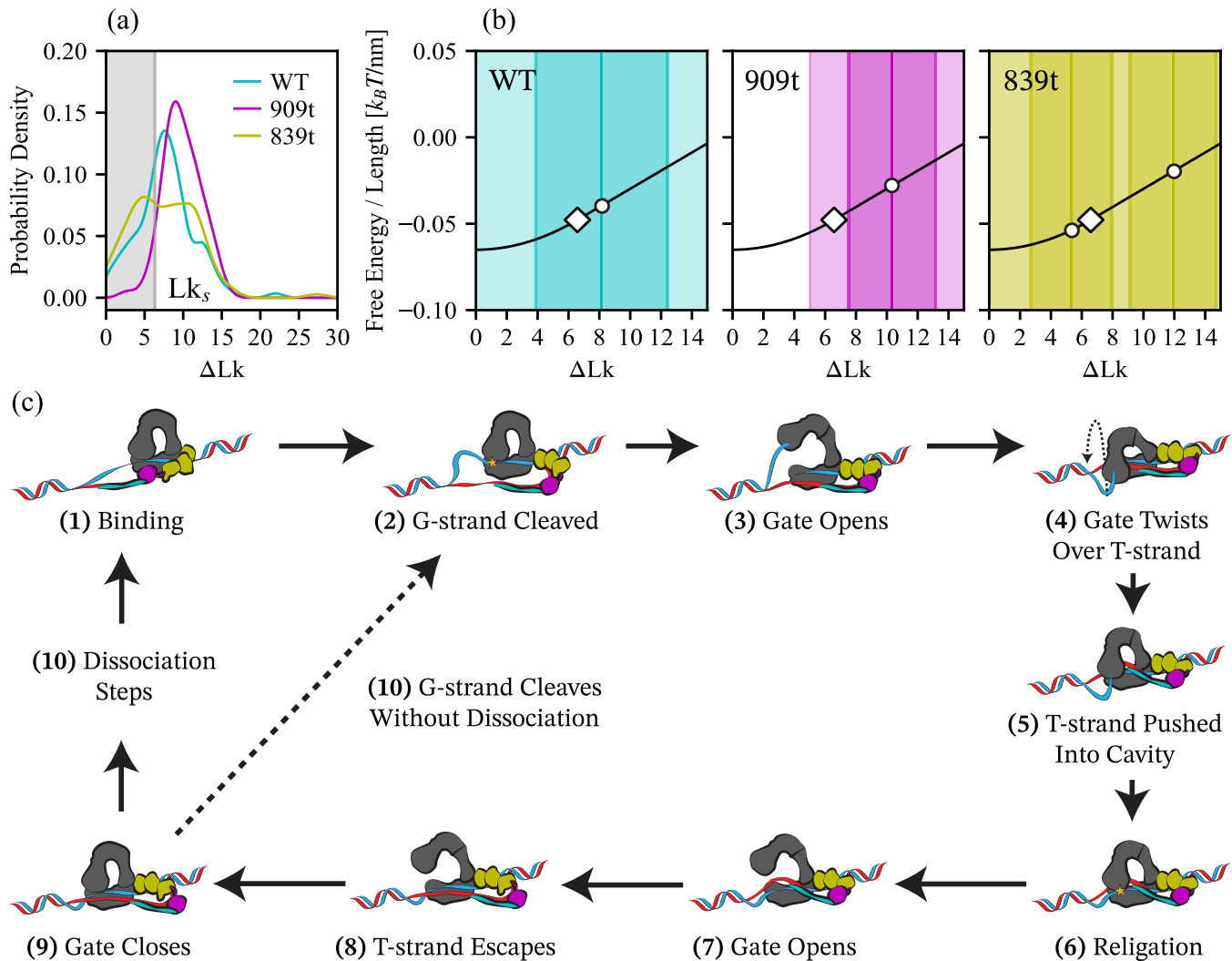


FIG. 7. Plots demonstrating the connection between the free energy profile of the DNA matching our experiment conditions, its buckling point, and how much linking number was remaining in the DNA before enzyme activity ceased. (a) Smooth histograms of the enzymes' recorded stopping positions. The dashed gray line represents the DNA buckling transition $Lk_s = 6.38 \Delta Lk$. ΔLk 's in the shaded region have only twist and no writhe, and ΔLk 's outside the shaded region have both twist and writhe. (b) Overlays of each protein's distribution from (a) on the free energy profile of the DNA. The solid black curves represent the DNA's free energy as a function of Lk , with a white diamond marking the DNA buckling transition. The solid vertical lines are the peaks of the Gaussians which fit the data, and a white circle marks where those distribution averages cross the energy curves. Regions are shaded dark within the full width half maximums of each Gaussian, and lightly shaded for the remaining data. (c) Our proposed strand-passage cycle which implements DNA corraling.

free energy per length. To prevent confusion in data presentation, plots exhibiting these parameters have been converted to linking number (ΔLk) instead of supercoiling density (σ) when possible; see Fig. 7. Changing them to use the linking number allows for direct comparison to our experimental extension curves (Fig. 4). It is important to note that the precision of measured extension is heavily lessened once a tether has relaxed to the buckling transition (Lk_s) due to the flattening of extension measurements for ΔLk smaller than Lk_s . However, it is still possible to tell that a tether has reached that threshold. After determining the parameter values for our system to use in the equations we overlaid the distributions of the proteins' stopping positions to examine their relationship with the DNA's energetic profile [Fig. 7(a)].

The results show the loss of writhe plays a role in the stopping point for the topoisomerases. The distributions for the stopping positions were fit with a multi-Gaussian fitting algorithm and the resulting Gaussian parameters were extracted. Parameters were used to create a comparison between the fits and free energy curves of the DNA used. The peak of the WT's distribution is $\sim 2 \Delta Lk$ before Lk_s [Fig. 7(b)], and the peak of the 909t's distribution $\sim 4 \Delta Lk$. In fact, the full width at half maximum (FWHM) of the 909t excludes the buckling transition entirely [Fig. 7(b)]. The same is true for the first peak and FWHM of the 839t, which is further behind the 909t [Fig. 7(b)]. In general, these energy differences highlight that the enzymes' functions are not just stochastically dependent on the DNA energy, but that the enzyme's ability to use the energy suffers with structural change. We hypothesize that

the most likely way the protein loses access to this energy is through increased mobility of the T strand during passage caused by loss of the CTDs. This also provides some additional context for the results of our 11 kb experiments. When comparing the energies for the 11 kb tethers at 0.65 pN, predicted by the model, there is consistently more energy available contained in the DNA up to the buckling point (Fig. S3 [33]). The higher energy may explain the increased activity we measured. This loss of usable energy combined with the results of Sec. III A hint at ways the cycle of strand passage occurs.

Writhe being removed also means there's less single stranded DNA. It has been shown by Cao and colleagues that the length of available DNA for the enzymes to bind to plays a role in cleavage [19]. Specifically, on shorter single stranded segments the CTDs may compete with the WT N-terminal core for binding, resulting in decreased cleavage and, by extension, relaxation rates. Unlike the WT, the 839t mutant showed increased cleavage on shorter single stranded segments, which could explain the second peak observed for the 839t beyond the buckling transition (Fig. 7). Given smaller amounts of denaturation as supercoils are removed, it follows that the 839t would have an easier time binding to the regions which remained while the WT would not.

C. Proposed strand-passage model

In general, it is accepted that the purpose of the C terminus is to aid in passing the T strand to the cavity, changing the DNA's linking number [18,19,23]. Our results provide insights about the CTDs effects on T strand passage. We propose an updated model which incorporates the information from our experiments. Previous works by many other groups have provided a foundation to build upon. What is presented here makes an effort to coalesce those works with our own data. As a result, in determining possible steps of the process, several concepts were assumed to remain true. First, the CTDs bind to the T strand and then undergo a conformational change which moves them away from the gate [Figs. 7(c1)–7(c5)]. Second, strand passage requires multiple attempts from the protein before it succeeds [39]. Third, the protein may stay bound to the substrate after successful strand passage. Fourth, the gate of the N-terminal core opens as described previously [11]. Fifth, the tail of the WT is bound adjacent to the other CTDs and there is enough room to accommodate it completely. Finally, we assume the CTDs do not enter the cavity.

A central goal of this work was determining the holistic effects of removing CTDs from the topoisomerase. To that end, we based all additions to the cycle on our experimental results. The outcome of this application of our data led to only two primary additions to the strand-passage cycle. Taking our results and the previously stated assumptions, we propose that strand passage occurs by a DNA corraling method [Fig. 7(c)]. In this model, when the CTDs move away the T strand is pulled tight, bringing it closer to the gate (2). After or while this occurs, the G strand is cleaved, and the gate opens (3). As D3 and D2 move up and away from D1, they twist toward the T strand. This twisting is driven by a mix of thermodynamics and a restoring torque in the cleaved strand. Once D3 has twisted past the T strand, the gate closes, trapping the T

strand between D1 and D3 (4). Given the fundamental process must remain the same, and that the CTDs are considered responsible for aiding passage of the T strand, we interpret our decreased run sizes as CTDs are removed as decreasing the capture rate of the T strand during this motion. This motion satisfies assumptions 2, 4, 5, and 1 by adding no additional conformational changes of the CTDs which fundamentally alter the strand passage, only its success rate. This could be due to changes in the DNA's interaction with the CTDs discussed in the next paragraph. Then, D3 and D2 begin to twist back toward D1, pushing, and ultimately locking, the T strand into the gate (5). The corraling of the T strand into the gate tightens the CTDs, forcing them to curl in toward the N-terminal core, and creates tension in the T strand. The cleaved strand is religated (6) and D3 detaches from the G strand, allowing the gate to be opened again (7). Seeking to relieve the stress from the curled CTDs, the T strand exits the cavity and straightens (8), allowing the domains to move away from the N-terminal core again (9). At this point, the protein may either dissociate or complete the cycle again (10).

As previously stated, we attempted to construct this cycle such that when CTDs are deleted, the overall cycle stays the same, but its properties could agree with our results. Each deletion of the CTDs effectively changes the enzyme's length, charge density, and binding affinity. The combined length of the CTDs, including the tail, could reduce the amount of T strand which is available for strand passage, possibly providing better control of the amount of supercoils removed. This could be important *in vivo* because DNA is typically in a supercoiled state due to its compaction. Further, the positive charge of the tail could reduce the negatively charged DNA's self-repulsion, allowing the strands to be brought closer to each other, aiding strand passage [Figs. 7(c4) and 7(c5)]. In other words, the changing run sizes of the studied proteins may be explained by difficulty trapping the T strand. This also provides a possible explanation for the reversal seen in run rates when compared to run sizes, i.e., the 909t having slower runs but resolving more turns. Others have shown separating the N- and C-terminal domains and complimenting them in a relaxation assay partially restores activity, but the N-terminal domain is unable to relax supercoils by itself [23]. This indicates that the electrostatic and structural changes to the T strand imparted by the CTDs are critical for strand passage and agrees with our results.

We posit the enzyme's failed attempts at strand passage appear as the decreases in the enzyme's run rate and the increases in dwells, both frequency and duration, within our data between variants. Our data show that, when compared to the WT enzyme, the 909t and 839t variants tend to dwell longer as more turns are relaxed on both sets of DNA and at either force [Figs. 6(a) and 6(b)]. Also, the 909t and 839t variants exhibit dwindling relaxation ability as they approach Lk_0 , and we attribute this to the flattening of the DNA's free energy in that regime (Figs. 6(a) and 6(b), and Fig. S3 in the Supplemental Material [33]). Run rate is being included because single relaxation steps are below our resolution limits; so frequent small dwells during bursts would be missed. This qualitatively explains the decrease in run rate from the WT to the 909t. Without the tail, the 909t may need more attempts at corraling the T strand before successfully capturing it. One

possibility for this behavior is that the tail stabilizes the T strand's position relative to the G strand, and without it the T strand cannot be manipulated as easily. Another is that the T strand may not be able to consistently be brought close enough to the gate to be passed. Considering the activity of the 839t, we believe it to most likely be the latter or a combination of both. Provided that the CTDs move away from the gate, losing D8 may force the T strand to be brought closer to the gate. This decrease in distance may allow the T strand to pass in fewer attempts than even the WT. However, strand passage would still be hindered by the lack of structure and maneuverability in the T strand, resulting in more pauses and shorter bursts of activity. In other words, it is a trade of consistency for speed. One could expect that as more CTDs are removed, the T strand distance becomes shorter, but control over it becomes lesser.

ACKNOWLEDGMENTS

We would like to thank Keir Neuman and Ian Morgan (NIH, NHLBI) for sharing PYTHON code to assist with data analysis; Connor Nance, Dayton Simpkins, Pavel Morales, Dylan Weaver, and Kara Balthrop for numerous long and helpful discussions; and Samuel Cheslik and Doniven Hicks for help with flow cell assembly. This work was supported by laboratory start-up funds from the University of Missouri and the National Institute of General Medical Sciences of the National Institutes of Health under Award No. R35GM156306.

DATA AVAILABILITY

The data that support the findings of this article are publicly available on Zenodo [40]. The program used for the analysis of data traces is also publicly available on Zenodo [41].

- [1] R. N. Irobalieva, J. M. Fogg, D. J. Catanese, T. Sutthibutpong, M. Chen, A. K. Barker, S. J. Ludtke, S. A. Harris, M. F. Schmid, W. Chiu, and L. Zechiedrich, Structural diversity of supercoiled DNA, *Nat. Commun.* **6**, 8440 (2015).
- [2] J. F. Marko, Torque and dynamics of linking number relaxation in stretched supercoiled DNA, *Phys. Rev. E* **76**, 021926 (2007).
- [3] J. C. Wang, L. J. Peck, and K. Becherer, DNA supercoiling and its effects on DNA structure and function, *Cold Spring Harbor Symp. Quant. Biol.* **47**, 85 (1983).
- [4] P. Bettotti, V. Visone, L. Lunelli, G. Perugino, M. Ciaramella, and A. Valenti, Structure and properties of DNA molecules over the full range of biologically relevant supercoiling states, *Sci. Rep.* **8**, 6163 (2018).
- [5] A. S. Krasilnikov, A. Podtezhnikov, A. Vologodskii, and S. M. Mirkin, Large-scale effects of transcriptional DNA supercoiling *in vivo*, *J. Mol. Biol.* **292**, 1149 (1999).
- [6] A. L. Mikheikin, A. Y. Lushnikov, and Y. L. Lyubchenko, Effect of DNA supercoiling on the geometry of Holliday junctions, *Biochemistry* **45**, 12998 (2006).
- [7] L. A. Freeman and W. T. Garrard, DNA supercoiling in chromatin structure and gene expression, *Crit. Rev. Eukaryotic Gene Expression* **2**, 165 (1992).
- [8] J. J. Champoux, DNA topoisomerases: Structure, function, and mechanism, *Annu. Rev. Biochem.* **70**, 369 (2001).
- [9] A. J. Schoeffler and J. M. Berger, DNA topoisomerases: Harnessing and constraining energy to govern chromosome topology, *Q. Rev. Biophys.* **41**, 41 (2008).
- [10] P. Forterre, S. Gribaldo, D. Gadelle, and M.-C. Serre, Origin and evolution of DNA topoisomerases, *Biochimie* **89**, 427 (2007).
- [11] M. Mills, Y.-C. Tse-Dinh, and K. C. Neuman, Direct observation of topoisomerase IA gate dynamics, *Nat. Struct. Mol. Biol.* **25**, 1111 (2018).
- [12] Z. Li, A. Mondragón, and R. J. DiGate, The mechanism of type IA topoisomerase-mediated DNA topological transformations, *Mol. Cell* **7**, 301 (2001).
- [13] N. H. Dekker, V. V. Rybenkov, M. Duguet, N. J. Crisona, N. R. Cozzarelli, D. Bensimon, and V. Croquette, The mechanism of type IA topoisomerases, *Proc. Natl. Acad. Sci. USA* **99**, 12126 (2002).
- [14] J. F. Marko, DNA under high tension: Overstretching, undertwisting, and relaxation dynamics, *Phys. Rev. E* **57**, 2134 (1998).
- [15] J. C. Wang, DNA topoisomerases, *Annu. Rev. Biochem.* **65**, 635 (1996).
- [16] T. Viard and C. B. de la Tour, Type IA topoisomerases: A simple puzzle? *Biochimie* **89**, 456 (2007).
- [17] K. D. Corbett and J. M. Berger, Structure, molecular mechanisms, and evolutionary relationships in DNA topoisomerases, *Annu. Rev. Biophys.* **33**, 95 (2004).
- [18] W. Ahmed, A. G. Bhat, M. N. Leelaram, S. Menon, and V. Nagaraja, Carboxyl terminal domain basic amino acids of mycobacterial topoisomerase I bind DNA to promote strand passage, *Nucl. Acids Res.* **41**, 7462 (2013).
- [19] N. Cao, K. Tan, X. Zuo, T. Annamalai, and Y.-C. Tse-Dinh, Mechanistic insights from structure of *Mycobacterium smegmatis* topoisomerase I with ssDNA bound to both N- and C-terminal domains, *Nucl. Acids Res.* **48**, 4448 (2020).
- [20] S. Ferdous, T. Dasgupta, T. Annamalai, K. Tan, and Y.-C. Tse-Dinh, The interaction between transport-segment DNA and topoisomerase IA—crystal structure of MtbTOP1 in complex with both G- and T-segments, *Nuc. Acids Res.* **51**, 349 (2023).
- [21] A. Seddek, T. Annamalai, and Y.-C. Tse-Dinh, Type IA topoisomerases as targets for infectious disease treatments, *Microorganisms* **9**, 86 (2021).
- [22] K. Buzun, A. Bielawska, K. Bielawski, and A. Gornowicz, DNA topoisomerases as molecular targets for anticancer drugs, *J. Enzyme Inhib. Med. Chem.* **35**, 1781 (2020).
- [23] P. Jain and V. Nagaraja, Indispensable, functionally complementing N and C-terminal domains constitute site-specific topoisomerase I, *J. Mol. Biol.* **357**, 1409 (2006).
- [24] I. De Vlaminck and C. Dekker, Recent advances in magnetic tweezers, *Annu. Rev. Biophys.* **41**, 453 (2012).
- [25] K. Kim and O. A. Saleh, A high-resolution magnetic tweezer for single-molecule measurements, *Nucl. Acids Res.* **37**, e136 (2009).
- [26] S. Gaire, R. Fabian, Jr., I. Pegg, and A. Sarkar, Magnetic tweezers: Development and use in single-molecule research, *BioTechniques* **72**, 65 (2022).

- [27] K. C. Neuman and A. Nagy, Single-molecule force spectroscopy: Optical tweezers, magnetic tweezers and atomic force microscopy, *Nat. Methods* **5**, 491 (2008).
- [28] Y. Seol and K. C. Neuman, Magnetic tweezers for single-molecule manipulation, in *Single Molecule Analysis: Methods and Protocols*, Methods in Molecular Biology, edited by E. J. G. Peterman and G. J. L. Wuite (Humana Press, Totowa, NJ, 2011), pp. 265–293.
- [29] D. Spakman, J. A. M. Bakx, A. S. Biebricher, E. J. G. Peterman, G. J. L. Wuite, and G. A. King, Unravelling the mechanisms of type IA topoisomerases using single-molecule approaches, *Nucl. Acids Res.* **49**, 5470 (2021).
- [30] F. Mosconi, J. F. Allemand, D. Bensimon, and V. Croquette, Measurement of the torque on a single stretched and twisted DNA using magnetic tweezers, *Phys. Rev. Lett.* **102**, 078301 (2009).
- [31] A. Celedon, I. M. Nodelman, B. Wildt, R. Dewan, P. Searson, D. Wirtz, G. D. Bowman, and S. X. Sun, Magnetic tweezers measurement of single molecule torque, *Nano Lett.* **9**, 1720 (2009).
- [32] M. J. Szafran, T. Strick, A. Strzalka, J. Zakrzewska-Czerwińska, and D. Jakimowicz, A highly processive topoisomerase I: Studies at the single-molecule level, *Nucl. Acids Res.* **42**, 7935 (2014).
- [33] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/9tkx-w3zj> for additional figures.
- [34] C. Mao and M. Mills, Characterization of human XPD helicase activity with single-molecule magnetic tweezers, *Biophys. J.* **123**, 260 (2024).
- [35] X. Gao, Y. Hong, F. Ye, J. T. Inman, and M. D. Wang, Torsional stiffness of extended and plectonemic DNA, *Phys. Rev. Lett.* **127**, 028101 (2021).
- [36] K. Terekhova, K. H. Gunn, J. F. Marko, and A. Mondragón, Bacterial topoisomerase I and topoisomerase III relax supercoiled DNA via distinct pathways, *Nucl. Acids Res.* **40**, 10432 (2012).
- [37] D. A. Koster, V. Croquette, C. Dekker, S. Shuman, and N. H. Dekker, Friction and torque govern the relaxation of DNA supercoils by eukaryotic topoisomerase IB, *Nature (London)* **434**, 671 (2005).
- [38] J. J. Champoux, Type IA DNA topoisomerases: Strictly one step at a time, *Proc. Natl. Acad. Sci. USA* **99**, 11998 (2002).
- [39] K. H. Gunn, J. F. Marko, and A. Mondragón, An orthogonal single-molecule experiment reveals multiple-attempt dynamics of type IA topoisomerases, *Nat. Struct. Mol. Biol.* **24**, 484 (2017).
- [40] D. Balthrop, D. Sigdel, C. Mao, Y. C. Tse-Dinh, and M. Mills, Data traces for supercoil relaxation of mycobacterial type IA topoisomerase and deletion variants [Data set], Zenodo, 2026, <https://doi.org/10.5281/zenodo.20671367>.
- [41] D. Balthrop, D. Sigdel, and M. Mills, Step analysis program for time series data (v1.0), Zenodo, 2026, <https://doi.org/10.5281/zenodo.20706183>.