Diffusion of knots in nanochannel-confined DNA molecules

Runfang Mao and Kevin D. Dorfman*

*Corresponding author. Email: dorfman@umn.edu

Department of Chemical Engineering and Materials Science, University of Minnesota - Twin Cities, 421 Washington Ave. SE, Minneapolis, Minnesota 55455, United States

Abstract

We used Langevin dynamics simulations without hydrodynamic interactions to probe knot diffusion mechanisms and the time scales governing the evolution and the spontaneous untying of trefoil knots in nanochannel-confined DNA molecules in the extended de Gennes regime. Knot untying follows an "opening up process", wherein the initially tight knot continues growing and fluctuating in size as it moves towards the end of the DNA molecule before its annihilation at the chain end. The mean knot size increases significantly and sub-linearly with the increasing chain contour length. The knot diffusion in nanochannel-confined DNA molecules is subdiffusive, with the unknotting time scaling with chain contour length with an exponent of 2.64 ± 0.23 to within 95% confidence interval. The scaling exponent for the mean unknotting time versus chain contour length, along with visual inspection of the knot conformations, suggests that the knot diffusion mechanism is a combination of self-reptation and knot region breathing for the simulated parameters.

1. Introduction

Knotting of DNA is a fundamental phenomenon, playing a key role in topological problems, polymer physics and biology.[1–4] For instance, the dynamics of knots within crowded environments *in vivo* is important for DNA replication,[5,6] cell survival[7] and cellular malfunction.[8] The presence of DNA knots also has a deleterious effect on the accuracy of genomics technologies, such as genome mapping in nanochannels and nanopore sequencing.[9,10] For example, in linear DNA, where only pseudo knots are possible, the knot diffusion time controls the rate of knot destruction and impacts the aforementioned genomic methods. Thus, it is important to understand how knots diffuse along the DNA contour. However, it is challenging to study DNA knot dynamics in free solution. It is both difficult to produce large numbers of knots and then detect and track knot motion along randomly-coiled single molecules in three dimensions. DNA knots can be efficiently generated using optical tweezers or an extensional flow field instead,[11–16] and these methods effectively reduce the knot tracking problem from a three-dimensional scenario to a quasi-one dimensional one. However, such technologies require stretching the DNA molecules by applying an external tension, which may confound the overall dynamics of knots when compared to the free solution case.

Nanochannel confinement provides an alternate way to study the dynamics of knotted DNA molecules.[17–20] Confinement has a unique advantage compared to optical tweezers or extensional flow since it generates knots by compression, leading to relaxed knotted DNA molecules that should be more similar to those formed in free polymers. Remarkably, even the

basic mechanism of knotted DNA in confinement remains a relatively open question. One possible diffusion mechanism is self-reptation, where the polymer contour follows a snake-like motion through the knot. Self-reptation is anticipated to be a subdiffusive process,[21,22] owing to the analogy with polymer translocation through a nanopore, and theory predicts that the diffusion time for self-reptation of a knot scales with contour length L as $\tau_D \sim L^3$ in the limit where the amount of polymer in the knot is a small fraction of the contour length.[23] Another possible diffusion mechanism is knot region breathing. In this case, the knotted region within the DNA molecule locally exchanges positions with its neighbors. Knot region breathing is anticipated to produce regular diffusion, and the diffusion time of knot region breathing gives a scaling $\tau_D \sim L^2$, again in the large L limit.[23]

To distinguish between these two possible mechanisms, Ma and Dorfman examined the diffusion of knots along relaxed DNA in nanochannels using a nanofluidic "knot factory" technique[18] to efficiently generate the knots in their experiments. The resulting measurement of the knot mean-squared-displacements indicated that the knot diffusion mechanism is likely to be self-reptation owing to the subdiffusive behavior.[17] However, their experiments do not constitute a direct test of the scaling theories for these two mechanisms. A more direct way to distinguish between the two diffusion mechanisms is to measure knot diffusion as a function of the degree of polymerization to identify the scaling exponent, or to directly compare the self-diffusion time with existing theories.[23] While it is challenging to adopt this approach experimentally due to the wide range of DNA molecular weights that are needed, it is relatively straightforward to vary L in simulations. Yet, simulation and experimental works on nanochannel-confined knots that have been performed mainly focus on establishing equilibrium properties, such as the knot formation probability and knot complexity.[18,24–30] Little work has been done focusing on the dynamic aspects of knot diffusion in

confinement.[31,32]

Here, we use molecular dynamics (MD) simulations of a coarse-grained polymer model to study the diffusion mechanism and the dynamics of 3_1 trefoil knots in nanochannel-confined DNA molecules. We track the knot size and knot position along the DNA molecule as a function of time. Additionally, we examine the impact of chain contour length L on both the mean knot size, as well as two dynamic properties, the knot mean square displacement (MSD) and the average unknotting time. Notably, the knot size increases with increasing L and the knot motion is subdiffusive when it is confined in a nanochannel. By comparing the diffusion dynamics and scaling exponent obtained from simulation and theories,[23] the simulations reveal that the knot diffusion is a combination of both self-reptation and knot region breathing mechanisms.

2. Methods

2.1 Pair potential models

The pairwise interaction between beads is modeled by purely repulsive Weeks-Chandler-Andersen (WCA) potential[33]

$$U_{\mathrm{WCA}} = \begin{cases} 4\varepsilon\left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right] + \varepsilon, r \leq 2^{\frac{1}{6}}\sigma \\ 0, r > 2^{\frac{1}{6}}\sigma \end{cases} \quad (1)$$

where $\varepsilon = k_{\mathrm{B}}T$. The bond interaction between adjacent beads is modeled by finite extensible nonlinear elastic (FENE) potential[34]

$$U_{\mathrm{FENE}} = -0.5k_{\mathrm{bond}}R_0^2 \ln\left[1 - \left(\frac{r_{ij}}{R_0}\right)^2\right], \; r_{ij} \leq R_0 \quad (2)$$

where the maximum bond length is $R_0 = 1.5\sigma$ and the bond stiffness is $k_{\mathrm{bond}} = 30\ k_{\mathrm{B}}T$. Chain

stiffness is implemented by a bond angle potential

$$U_{\text{angle}} = 0.5\kappa(\theta_i - \pi)^2 \quad (3)$$

where the bending penalty κ = 5 k_BT is obtained by mapping to the persistence length l_p of DNA[19] through the relation with l_p = 60 nm and the bead diameter $w = \sigma$ = 12 nm,

$$\frac{l_p}{w} = \frac{\kappa}{\kappa - \kappa\coth(\kappa) + 1} \quad (4)$$

The polymer chain is confined in nanochannel along the *y*- and *z*-axis with a wall potential described by WCA potential

$$U_{\text{wall}} = \begin{cases} 4\varepsilon\left[\left(\frac{\sigma}{d_i}\right)^{12} - \left(\frac{\sigma}{d_i}\right)^{6}\right] + \varepsilon, d_i \le 2^{\frac{1}{6}}\sigma \\ 0, d_i > 2^{\frac{1}{6}}\sigma \end{cases} \quad (5)$$

where d_i is the orthogonal distance between bead *i* and wall.

2.2 Initial configuration generation

The DNA molecule is modeled as a string of beads connected by the FENE bonds described in section 2.1. To generate the initial configuration of the knotted DNA, the two parts of an initial configuration were generated separately. The first part is the relaxed polymer chain. To generate this part of the initial configuration, we used molecular dynamics (MD) simulations, described in more detail in section 2.3, to equilibrate a DNA strand from an initially fully stretched DNA state within the channel for a total 5×10^8 time steps for each different chain contour length. About 1×10^8 time steps are required for DNA molecules to fully relax. Uncorrelated configurations of DNA molecules were then selected from the remaining 4×10^8 time steps, with a sampling interval based on autocorrelation time (SI **Figure S1**). For the longest chain *L* = 600, which shows slower correlation time, one additional parallel simulation with a different random seed was performed to collect sufficient uncorrelated samples. The

second part of the initialization produces the relatively tight 3_1 trefoil knot configuration, which we set to contain 14 beads. The initial configuration of the trefoil knot was generated by Knotplot[35] with a bead diameter of σ. The 3_1 trefoil knot was then inserted into the middle of a relaxed DNA molecule sampled from the first protocol to generate one initial configuration; inserting this knot into each relaxed DNA sample generates an ensemble of initial conditions with different configurations of the unknotted portion of the DNA and identical knots. This procedure was used to generate initial conditions with the chain contour length L = 100 to 600 in units of the equilibrium bond length. **Figure 1** shows three examples of initial configurations of trefoil knots in nanochannel-confined DNA molecules with different L. For each L, at least 30 different initial configurations were generated. The topology of 3_1 knots was confirmed by calculating the Alexander polynomial[36] to avoid artifacts from bad initial configurations.

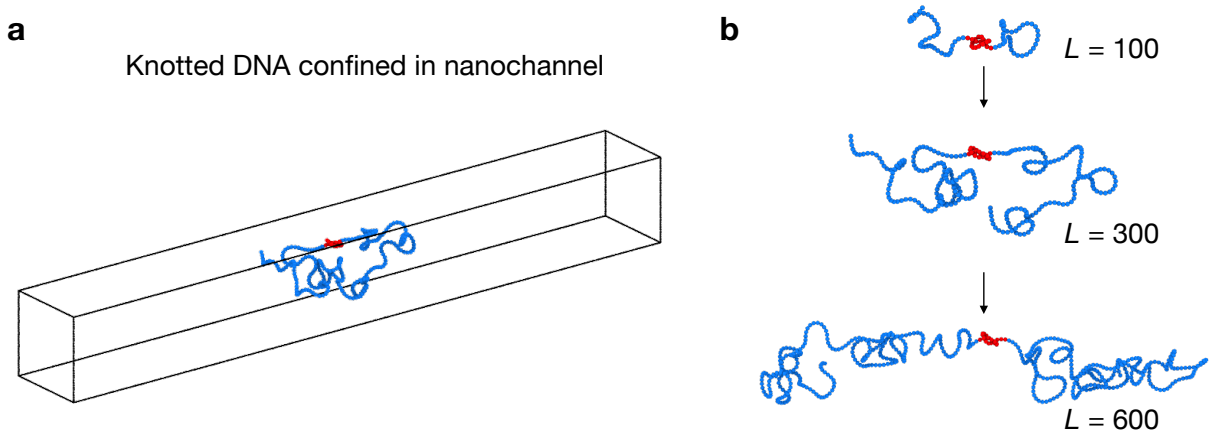

Figure 1. (a) Initial configurations of 3_1 trefoil knots in nanochannel-confined DNA molecules in the extended de Gennes regime. (b) Three knotted DNA molecules with different chain contour lengths by increasing the total number of beads L = 100, 300 and 600. The trefoil knots identified by Kymoknot are highlighted as red while the unknotted polymer parts are colored blue.

2.3 Molecular dynamics simulations

To study the diffusion of DNA knots in confinement, molecular dynamics (MD) simulations were performed using LAMMPS[37] in the canonical ensemble (NVT) using a

Langevin thermostat with a damping coefficient τ = 2. Note that the Langevin thermostat models an implicit solvent, so there are no hydrodynamic interactions in this model. Relaxed knotted DNA molecules were initially placed in the center of the simulation with a box size of $L \times 25 \times 25$, where L is the chain contour length. The geometries of knotted DNA molecules were subject to local energy minimization before being used for MD simulations. The periodic boundary condition in x-dimension was applied while the wall potential on y- and z- dimensions were applied to enforce confinement. The channel size D = 25 σ is mapped to experiment from Ma and Dorfman,[17,19] corresponding to a channel size of 300 nm in the extended de Gennes regime. Each simulation was conducted with total 1×10^8 time steps for production with step size $\Delta t = 0.001\ \tau_{MD}$, where $\tau_{MD} = \sigma(m/\varepsilon)^{1/2}$ is the Lennard-Jones (LJ) time based on the bead mass m. The first 80% of the simulation data of the knot size versus the last 80% of simulation data were compared to make sure no additional equilibrium time is required after local energy minimization. Simulations were performed isothermally at reduced temperature $T^* = 1$. All quantities are presented in LJ reduced units, and all particle beads have equal masses. Simulation configurations were dumped every 1×10^4 time steps for further analysis.

2.4 Structural analysis

The knotted topology on DNA molecules was characterized by calculating the Alexander polynomial of the knot using Kymoknot.[38] For linear chains, the knot topology can be determined by bridging the two terminal ends of the DNA molecules. The knot topology from a pseudo-closed ring, for which the topological state is well-defined, can thus be identified. Several such closing techniques have been previously proposed.[38–40] Here, we adopted a bottom-up procedure that can identify the knot topology on a linear chain, where the knot is searched staring from very short portions of the molecule and gradually considering longer ones. The knot is localized until the subchain used for the knot calculation contains a physical

knot of the same type as the whole chain, and the remaining portions are physically unknotted.[38] The knot start and end bead indices can then be identified based on the bottom-up knot search approach. The size of knotted DNA molecules was further quantified by the total number of beads contained in the knot, N_{knot}, which was calculated as the difference between the identified knot start and end bead indices.

2.5 Mean square displacement (MSD)

The locations of the edges of a knot were computed by using the Alexander polynomial through the bottom-up approach discussed in Section 2.4. The center of mass of the DNA knot was then calculated as

$$x_{\text{com}} = \frac{1}{M_{\text{knot}}} \sum_{i=1}^{N_{\text{knot}}} m x_i \tag{6}$$

where M_{knot} is the total mass of all beads contained within the knot and N_{knot} is the number of beads contained within the knot. All beads that are contained in the knot have equal mass m = 1 and are located along the channel axis with coordinates x_{i}, i = 1, …, n.

The time evolution of the center of mass, x_{com}, was used to compute the ensemble-averaged MSD

$$MSD(\delta t) = \langle [x_{\text{com}}(t) - x_{\text{com}}(t - \delta t)]^2 \rangle_{\text{t,n}} \tag{7}$$

where $\langle \cdots \rangle_{t,n}$ denotes an average over all times t and the ensemble of n parallel simulation trajectories, and δt is the time lag between the recorded simulation frames.

To analyze the diffusive behavior of knots, we computed the scaling exponent β of the ensemble-averaged MSD by fitting the logarithm of the data with a linear function[17]

$$\log_{10} MSD\,(\delta t) = \beta \log_{10} \delta t + c \tag{8}$$

where β and c are both fitted constants. The choice of upper bound used for fitting was

determined by calculating the correlation coefficient, R^2, of the linear fit with different choice of upper bounds and a fixed lower bound at $\delta t = 10\ \tau_{MD}$. The upper bound of time lag 50 τ_{MD} was determined based on the highest R^2 value obtained from all fittings.

2.6 Unknotting time

The unknotting time τ_u can be directly obtained from simulation, which is defined as the time when an initially knotted DNA molecule reaches an unknotted state for the first time. Since the unknotted chain can reform a knot through diffusion of the chain end, an unknotted state is defined as the absence of a knotted configuration from 10 continuous recorded simulation frames, i.e., over 10^5 time steps. To obtain insights into the knot diffusion mechanism, we computed the scaling exponent β of the ensemble-averaged τ_u by fitting the logarithm of the data with a linear function

$$\log_{10} \tau_u = \beta \log_{10} L_c + c \tag{9}$$

where β and c are both fitted constants.

Diffusion time models were proposed for both self-reptation and knot region breathing mechanisms by Metzler et al.[23] in the limit where the knot comprises a small part of the overall chain contour length. This theory will be compared with the results obtained from our MD simulations, so it is useful to review the key results here. The diffusivity for knot region breathing can be characterized by $D_{KRB} \simeq \alpha k_B T/\Delta$, where Δ is the length of the knot and α is the inverse of the bead friction, which can be obtained from a free polymer simulation. The knot diffusion time based on its initial location inside the chain (x_L) to one of the ends is given by[23]

$$\tau(x_L) = \int_0^{x_L} \frac{\frac{L_{red}}{2} - x'}{D(x')} dx' \tag{10}$$

where $L_{red} = L - \Delta$ is the reduced length of the two linear segments of the chain, L is the chain length and Δ is the knot length. To approximate the knot length, we used total number of beads contained in the knot, N_{knot}, which can be obtained from the simulation directly. The diffusion time for knot region breathing is given by[23]

$$\tau_{\mathrm{KRB}}(x_L) = \frac{\Delta L_{\mathrm{red}}^2}{2\alpha k_B T}\left[\left(\frac{x_L}{L_{\mathrm{red}}}\right) - \left(\frac{x_L}{L_{\mathrm{red}}}\right)^2\right] \quad (11)$$

In the self-reptation mechanism, the diffusivity is characterized by[23]

$$D_{SR} \simeq \frac{\alpha k_B T(L+\Delta)}{(x_L+\Delta)(L-x_L)} \quad (12)$$

From Eq. (10), the diffusion time of self-reptation is given by[23]

$$\tau_{\mathrm{SR}}(x_L) = \frac{1}{12\alpha k_B T(L+\Delta)}\left[x_L\{3LL_{\mathrm{red}} - 4Lx_L - 2L_{\mathrm{red}}x_L + 3x_L^2\right.$$

$$\left.+\Delta(6L[L_{\mathrm{red}} - x_L] + x_L[4x_L - 3L_{\mathrm{red}}])\}\right] \quad (13)$$

Note there was a typo "$3x_L^3$" in Eq. (6) in Ref. 23, which has been corrected in our Eq. (13).

3. Results

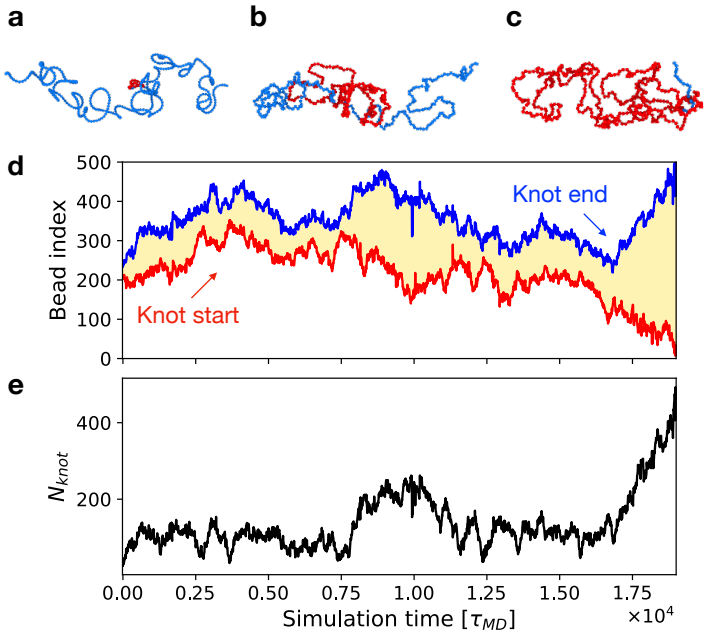

Figure 2. Snapshots for $L = 500$ at simulation time τ_{MD} of (a) 0, (b) 9.8×10^6 and (c) 1.9×10^7. The knotted region identified by Kymoknot is colored red while the unknotted polymer part is colored blue. (d) The start (red line) and end (blue line) of the bead indices contained in a 3_1 trefoil knot along DNA chain for the trajectory used to generate the snapshots in panels (a,b,c). (e) The number of beads in the knot, N_{knot}, as a function of simulation time calculated from (d).

To elucidate the dynamics of spontaneous knot untying of linear DNA molecules in confinement, we monitored both the knot position and the knot size, characterized by the number of beads in the knot, N_{knot}, during the simulations. **Figure 2** shows the swelling of a trefoil knot as a function of time. The trefoil knot diffuses away from its original central position, fluctuating along the entire DNA molecule. During the simulation, the initially tight knot continues growing and fluctuating in size as it moves towards the end of the DNA. The knot is destroyed upon reaching the terminus of the DNA molecule. The phenomenology in **Figure 2** was observed across many simulations with this parameter set; corresponding results for different initial configurations are provided in SI **Figures S2 and S3**. In general, we found that knots open up, although the value of N_{knot} fluctuates between different simulations. Notably, we also found that the maximum N_{knot} is restricted by the total size of the chain and below the trivial limit of $N_{\text{knot}} = L$. As shown in **Figure S4**, N_{knot} reaches a maximum size of

around 150 at L = 300. Instead, at L = 500 (**Figure 2**), N_{knot} can reach a maximum of around 400. These observations suggest that the knot untying follows an "opening up process", and the fraction of the chain that participates in the opening up process is related to L.

This knot opening up process agrees qualitatively with previous single-molecule experiments,[16] wherein knot swelling was reported in linearly extended DNA molecules during relaxation, although the actual knot topology in experiments may be more complicated than the simple 3_1 trefoil knot studied here. The behavior in **Figure 2** also agrees with Monte Carlo simulations of long flexible linear polymers, suggesting a growth of average knot length on the total chain length.[41]

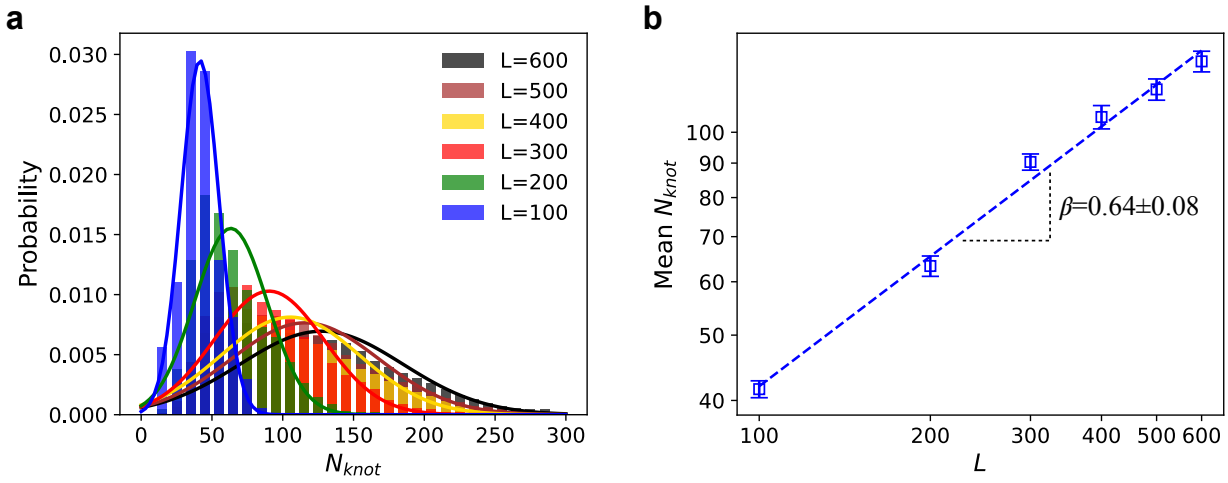

Figure 3. (a) Histogram of N_{knot} as a function of chain contour length L. (b) Mean N_{knot} as a function of L. The intercept c from fitting to a power law is 0.34. The error is estimated using a 95% confidence interval.

To elucidate the mechanism behind knot untying, it proves useful to quantify the knot size with respect to L. **Figure 3** shows the number of beads contained within the knot, N_{knot}, with respect to L. Since the unknotting dynamic is not an equilibrium process, N_{knot} is collected from the start of the simulation until the knot is fully untangled in all parallel simulations. The probability distributions of $N_{knot}(L)$ in **Figure 3a** can be well-fitted to a normal Gaussian

distribution. Aside from this phenomenological observation, two insights into the dependence of N_{knot} on L emerge by examining the probability distribution. First, the length-dependent increase in N_{knot} is clearly illustrated by the location of the peaks of the distributions. Secondly, the spread of the probability distribution also increases linearly with increasing N_{knot}, indicating the distributions are self-similar as shown in **Figure S5**. The sizable number of knot configurations during a simulation allows us to further compute the mean value and the statistical error of observables. **Figure 3b** shows the log-log plot of the calculated mean N_{knot} , where the true statistical error on the correlated data was computed from the block averaging method.[42] The length-dependent increase of N_{knot} fits a power law with exponent $\beta = 0.64 \pm 0.08$, where error refers to a 95% confidence interval from the linear regression.

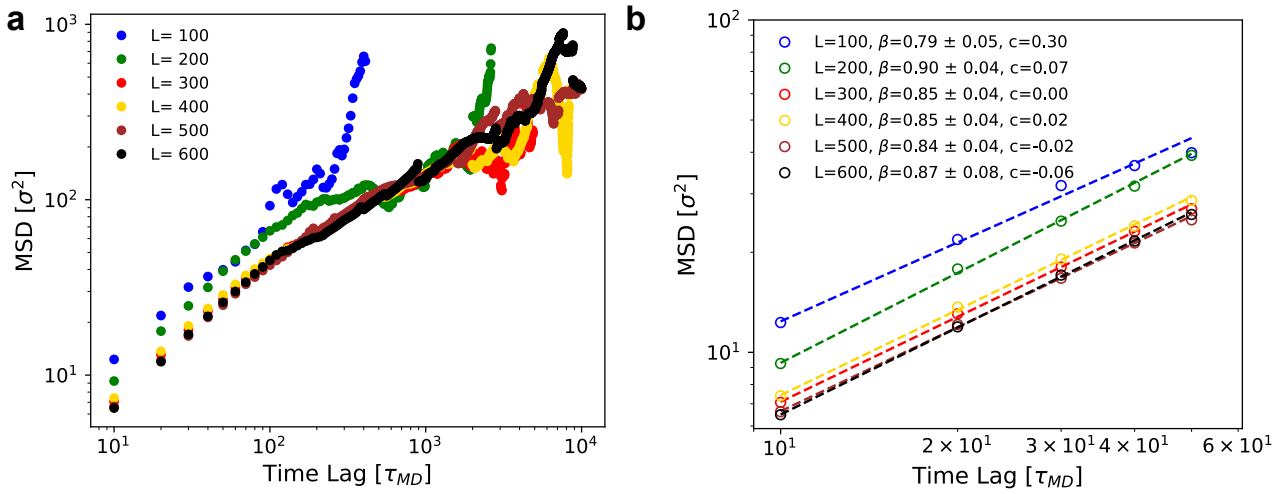

Figure 4. (a) Ensemble-averaged mean square displacements as a function of time lag for different L. The time-averaged MSDs for calculating ensemble-averaged MSD from each trajectory is provided in **Figure S6**. (b) Linear fit to the logarithm of ensemble-averaged MSD and time lag window between 10 to 50 τ_{MD}. The choice of time lag window is discussed in Section 2.5.

The evolution of the center-of-mass of knots in time further allows us to calculate the knot mean-square displacement. **Figure 4** shows the ensemble-averaged MSD as a function of time lag for different chain contour lengths. The knot diffusive behavior is further quantified by

fitting the logarithm of ensemble-averaged MSD and the time lag window between 10^4 to 5×10^4, which avoids the localization error (at short time lags) and the sampling error (at long time lags). The range of scaling exponent extracted from fitting in **Figure 4** is [0.79,0.90] despite the difference in L. Notably, there is no apparent correlation between the scaling exponent and chain contour length. One possible reason for not seeing correlation between the scaling exponent and L can be large statistical uncertainty due to the small dataset of knotted state. In any event, all calculated scaling exponents indicate subdiffusive motion of knots along DNA molecules confined in nanochannel at short times.

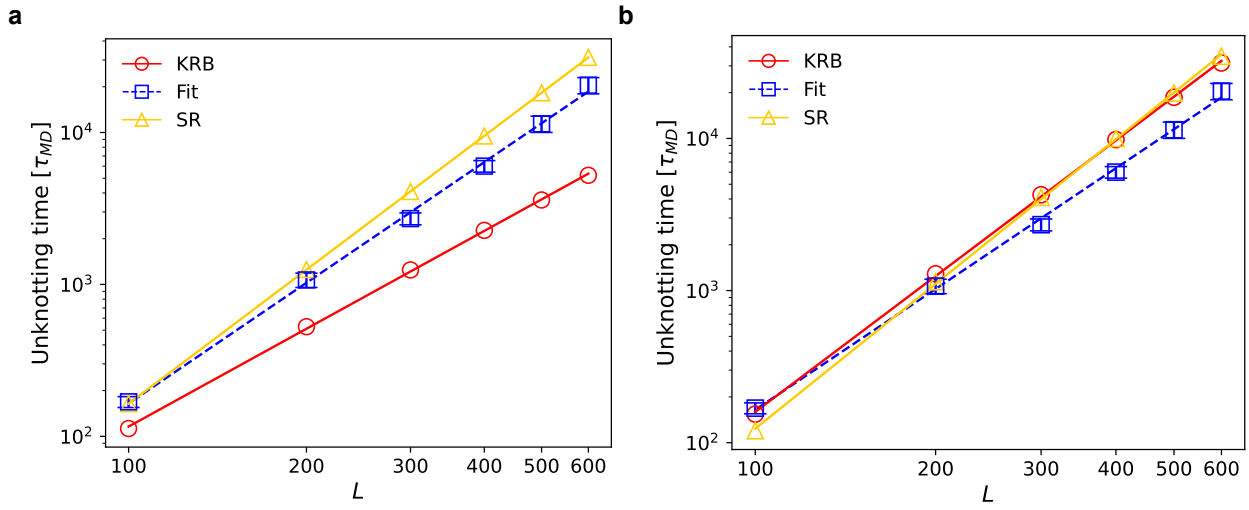

Figure 5. (a,b) Mean unknotting time (blue squares) obtained from the MD simulations and theories for knot region breathing (red circles) and self-reptation (gold triangles) as a function of L using (a) N_{knot} = 14 and (b) $N_{knot} = 2.19L^{0.64}$. The blue dashed line is the best fit to the logarithm of the mean unknotting time and chain contour length obtained from the MD simulations. The red and gold lines are the unknotting time predicted from the knot region breathing model and the self-reptation model, respectively.[23] (a) Using N_{knot} = 14, the slope and intercept predicted from the knot region breathing model are 2.14 and -2.22, respectively, while the slope and intercept predicted from the self-reptation model are 2.93 and -3.64, respectively. (b) Using $N_{knot} = 2.19L^{0.64}$, the slope and intercept predicted from the knot region breathing model are 2.96 and -3.73, respectively, while the slope and intercept predicted from the self-reptation model are 3.16 and -4.23, respectively. The error bars are the standard error of the mean.

The mean time required to spontaneously untie the trefoil knot is shown as blue squares in **Figure 5**. The mean unknotting time grows appreciably with L. These values can be well-fitted

298 to a power law with an exponent of 2.64 ± 0.23, where the error refers to a 95% confidence
299 interval from the linear regression. Similar observations have been reported for knots on
300 tension-free, open linear chains, where knots exhibit larger size and survival time for longer
301 chains despite the difference in model and simulation setup.[43] The magnitude of unknotting
302 time also agrees with the reported Brownian simulations of semiflexible open chains confined
303 in nanochannel (see supplementary material).[44]

304 The scaling exponent can be further compared with the existing theories proposed by
305 Meltzer et al., namely the self-reptation (SR) and knot region breathing (KRB) mechanisms.[23]
306 However, this comparison is not as straightforward as one might initially expect, since the
307 theory assumes that the knot is small compared to the total length of the chain, but **Figure 3**
308 shows that the knots are a significant fraction of their total chain contour and can grow
309 significantly (e.g., **Figure 2**), especially towards the end of the untying process.

310 Let us first consider the naïve case where we assume that $\Delta = N_{\mathrm{knot}}$ is small compared to L
311 by fixing the knot size to that of the initial tight trefoil knot, $N_{\mathrm{knot}} = 14$. **Figure 5a** compares
312 the predictions of Eqs. (11) and (13) using $N_{\mathrm{knot}} = 14$ to the simulation data. The apparent
313 scaling exponents from fitting the theory predictions over this limited range in L are 2.93 for
314 SR and 2.14 for KRB. For this choice of a small N_{knot}, the quantitative agreement between the
315 prediction of SR in Eq. (13) and the simulation is very good; this agreement is possible despite
316 the disagreement in scaling exponents because our simulations span less than one decade in L.
317 If we instead probe out to very large L for the theory (**Figure S7**), the simulation data are
318 approaching the predictions of the self-reptation model, suggesting that the difference in slope
319 would be rectified for sufficiently large L. Hence, if we take as an assumption that $N_{\mathrm{knot}} \ll L$,
320 the simulations support the self-reptation model.

321 However, the quantitative agreement between the simulation data and self-reptation in

Figure 5a, as well as the disagreement with knot region breathing, is an artifact of assuming a small knot. In the MD simulations, we found that the knot size is not much smaller than the total contour length, as suggested by the theory of Meltzer et al., but rather is a significant fraction of the total chain length (**Figure 3a**). Thus, we repeated the analysis of Melzer et al., replacing $N_{knot} = 14$ with the empirical results $N_{knot} = 2.19L^{0.64}$ regressed from **Figure 3b**. As shown in **Figure 5b**, the apparent diffusion time now scales as 3.16 for the self-reptation and 2.96 for the knot region breathing mechanism respectively. When taking into account that N_{knot}/L is not small (*i.e.*, sensible N_{knot}/L limit), it is no longer possible to distinguish between the different diffusion mechanisms and the simulation data; for the sensible N_{knot}/L limit, the scaling exponent 2.64 ± 0.23 extracted from the simulation does not agree perfectly well with the prediction either from KRB or SR theory. When taking in conjunction with our observations of the knot dynamics, the results suggest that knot diffusion is a combination of both mechanisms.

4. Discussion

The most straightforward method to characterize the unknotting mechanism is to analyze the nature of the diffusion itself. Anomalous subdiffusive dynamics are considered in many publications to be an indicator for the self-reptation mechanism, while the regular diffusive behavior is suggested for knot region breathing in the large L limit.[23] A previous study by Klotz et al. noted the anomalous diffusive behavior of knots in DNA molecules under an elongational field in microfluidic devices, where the knot motion is mediated by self-reptation.[15] Additionally, computational work by Matthews et al. reported a subdiffusive motion of knots along stretched polymers at short times, again illustrating that the knots diffuse through a mechanism similar to reptation.[45] In addition to these studies of the dynamics of knots under

external tensions, Ma and Dorfman[17] observed the subdiffusive motion of knots under nanochannel confinement, which again suggested a self-reptation mechanism. Our MSD results for different values of L (**Figure 4b**) clearly demonstrate that the knot motion is anomalous, with subdiffusive scaling exponents at short times under nanochannel confinement, consistent with all of the aforementioned studies.[15,17,23,45] However, our analysis suggests that even if the knot MSD shows a subdiffusive scaling, that does not imply strictly self-reptation behavior for the reasons noted in the discussion of **Figure 5**. This result suggests caution when inferring the knot diffusion mechanism solely from the MSD data.

There are several salient features of the knot dynamics that appear to arise due to confinement, in contrast to the previous literature on stretched polymers where the knots exhibit self-reptation along the entire molecule.[15,45] When the polymer is stretched, the knot remains localized and tight during its motion, only losing its tightness when it is destroyed when it diffuses to the end of molecule.[46,47] Furthermore, knots on tense chains can remain jammed with a fairly constant and tight knot size in a relatively long time period, and the knot diffusivity decreases exponentially with increasing tension.[47] Several qualitative differences were observed in knot dynamics between knots that are under tension and those under confinement. Firstly, in our MD simulations, we observed that the knot size fluctuates substantially during the entire process. Secondly, the knot sizes and knot positions change simultaneously and fluctuate frequently along the DNA molecules. As depicted in **Figure 2** and SI **Figures S2-4**, we found no sign of a jammed state in our simulations, i.e., where the knot remained stagnant with a relatively constant size over time. Finally, the knots follow an opening-up processes, with knots typically growing in size via power laws that depend on L at the short chain length limit between $L = 100$ to 600. There are thus several qualitative signatures that distinguish diffusion of confined knots and knots on stretched chains.

The direct observation of knot dynamics and comparison of the scaling exponent of knot diffusion time vs L with the existing theories[23] illustrates that both SR and KRB mechanisms contribute to knot diffusion at the sensible N_{knot}/L limit for nanochannel confinement. The significant difference in the prediction of apparent scaling exponents in the KRB at different N_{knot}/L limit also indicate that the knot size plays an important role in governing the diffusion time. Both mechanisms appear to be active on our simulations, and the results do not correspond directly to the limiting cases in the theory of Meltzer et al.[23] When the chain becomes longer (or confinement becomes stronger), the knot size should become increasingly small when compared to the total chain length, and it would be interesting to see if the results indeed approach the predictions of Meltzer et al.[23] However, longer chains and increased confinement may impact the geometry and size of the knot differently, leading to different scaling exponents and diffusion times. This is a promising direction for future work to develop a complete understanding of the diffusion of knots along DNA molecules.

5. Conclusions

In the present contribution, we explored the diffusion mechanism of trefoil knots with different chain contour lengths in confined systems using coarse-grained molecular dynamics simulations. The knot was observed to swell and continuously fluctuate in size during the simulation, in contrast to the self-reptation behavior that is reported for tight knots diffusing on stretched polymers. We observed subdiffusive motion of the knot, in agreement with the anomalous subdiffusive dynamics for the self-reptation mechanism. However, our observations of the knot structure indicate that, even if the knot MSD shows subdiffusive scaling, this does not strictly imply self-reptation behavior. Rather, the knot diffusion is a combination of both self-reptation and knot region breathing.

Supplementary Material

Supplementary information is available in the online version of the paper. Autocorrelation function of radius of gyration for different chain contour length L; Additional quantification of knotted DNA trajectories at L=300,500; The standard deviation in N_{knot} as a function of N_{knot}; details for calculating ensemble-averaged MSD from time-averaged MSD; Additional prediction of unknotting time vs. L.

Acknowledgements

This work was supported by National Science Foundation through award CBET-2016879. Computational resources were provided by the Minnesota Supercomputing Institute.

Author Declarations

Conflicts of interest

The authors have no conflicts to disclose.

Author Contributions

Runfang Mao: Data curation (lead); Investigation (lead); Methodology (lead); Formal analysis (equal); Software (lead); Validation (equal); Visualization (lead); Writing – original draft (lead). **Kevin D. Dorfman**: Conceptualization (lead); Formal analysis (equal); Validation (equal); Funding acquisition (lead); Project administration (lead); Resources (lead);

Supervision (lead); Writing – review & editing (lead).

Data Availability

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

References

[1] D. Meluzzi, D.E. Smith, and G. Arya, Annu. Rev. Biophys. **39**, 349 (2010).

[2] L.F. Liu, L. Perkocha, R. Calendar, and J.C. Wang, Proc. Natl. Acad. Sci. **78**, 5498 (1981).

[3] W.R. Taylor and K. Lin, Nature **421**, 25 (2003).

[4] E. Orlandini, J. Phys. A Math. Theor. **51**, 53001 (2017).

[5] D. O'Donnol, A. Stasiak, and D. Buck, Nucleic Acids Res. **46**, 9181 (2018).

[6] L. Olavarrieta, M.L. Martínez-Robles, P. Hernandez, D.B. Krimer, and J.B. Schvartzman, Mol. Microbiol. **46**, 699 (2002).

[7] S.H. Chen, N.-L. Chan, and T. Hsieh, Annu. Rev. Biochem. **82**, 139 (2013).

[8] K. Hevener, T.A. Verstak, K.E. Lutat, D.L. Riggsbee, and J.W. Mooney, Acta Pharm. Sin. B **8**, 844 (2018).

[9] E.T. Lam, A. Hastie, C. Lin, D. Ehrlich, S.K. Das, M.D. Austin, P. Deshpande, H. Cao, N. Nagarajan, M. Xiao, and others, Nat. Biotechnol. **30**, 771 (2012).

[10] M. Jain, H.E. Olsen, B. Paten, and M. Akeson, Genome Biol. **17**, 1 (2016).

436 [11] X.R. Bao, H.J. Lee, and S.R. Quake, Phys. Rev. Lett. **91**, 265506 (2003).

437 [12] B.W. Soh, A.R. Klotz, and P.S. Doyle, Macromolecules **51**, 9562 (2018).

438 [13] B.W. Soh, V. Narsimhan, A.R. Klotz, and P.S. Doyle, Soft Matter **14**, 1689 (2018).

439 [14] V. Narsimhan, A.R. Klotz, and P.S. Doyle, ACS Macro Lett. **6**, 1285 (2017).

440 [15] A.R. Klotz, B.W. Soh, and P.S. Doyle, Phys. Rev. Lett. **120**, 188003 (2018).

441 [16] A.R. Klotz, V. Narsimhan, B.W. Soh, and P.S. Doyle, Macromolecules **50**, 4074 (2017).

442 [17] Z. Ma and K.D. Dorfman, Macromolecules **53**, 6461 (2020).

443 [18] S. Amin, A. Khorshid, L. Zeng, P. Zimny, and W. Reisner, Nat. Commun. **9**, 1 (2018).

444 [19] Z. Ma and K.D. Dorfman, J. Chem. Phys. **155**, 154901 (2021).

445 [20] Z. Ma and K.D. Dorfman, Macromolecules **54**, 4211 (2021).

446 [21] P.G. Dommersnes, Y. Kantor, and M. Kardar, Phys. Rev. E **66**, 31802 (2002).

447 [22] A.Y. Grosberg and Y. Rabin, Phys. Rev. Lett. **99**, 217801 (2007).

448 [23] R. Metzler, W. Reisner, R. Riehn, R. Austin, J.O. Tegenfeldt, and I.M. Sokolov, EPL **76**,
449 696 (2006).

450 [24] L. Dai and P.S. Doyle, Macromolecules **51**, 6327 (2018).

451 [25] L. Dai, J.R.C. van der Maarel, and P.S. Doyle, ACS Macro Lett. **1**, 732 (2012).

452 [26] C. Micheletti and E. Orlandini, Macromolecules **45**, 2113 (2012).

453 [27] A. Jain and K.D. Dorfman, Biomicrofluidics **11**, 24117 (2017).

454 [28] C. Micheletti and E. Orlandini, Soft Matter **8**, 10959 (2012).

455 [29] E. Orlandini and C. Micheletti, J. Biol. Phys. **39**, 267 (2013).

456 [30] C. Plesa, D. Verschueren, S. Pud, J. Van Der Torre, J.W. Ruitenberg, M.J. Witteveen, M.P.

457 Jonsson, A.Y. Grosberg, Y. Rabin, and C. Dekker, Nat. Nanotechnol. **11**, 1093 (2016).

458 [31] A. Suma, E. Orlandini, and C. Micheletti, in *APS March Meet. Abstr.* (2017), pp. M1--262.

459 [32] A. Suma, E. Orlandini, and C. Micheletti, J. Phys. Condens. Matter **27**, 354102 (2015).

460 [33] J.D. Weeks, D. Chandler, and H.C. Andersen, J. Chem. Phys. (1971).

461 [34] G.S. Grest and K. Kremer, Phys. Rev. A **33**, 3628 (1986).

462 [35] R. Scharein, Progr. Draw. Vis. Manip. Energy Minimizing Knots. See Http//Www.
463 Knotplot. Com (1998).

464 [36] J.W. Alexander, Trans. Am. Math. Soc. **30**, 275 (1928).

465 [37] A.P. Thompson, H.M. Aktulga, R. Berger, D.S. Bolintineanu, W.M. Brown, P.S. Crozier,
466 P.J. in 't Veld, A. Kohlmeyer, S.G. Moore, T.D. Nguyen, R. Shan, M.J. Stevens, J.
467 Tranchida, C. Trott, and S.J. Plimpton, Comp. Phys. Comm. **271**, 108171 (2022).

468 [38] L. Tubiana, G. Polles, E. Orlandini, and C. Micheletti, Eur. Phys. J. E **41**, 1 (2018).

469 [39] L. Tubiana, E. Orlandini, and C. Micheletti, Prog. Theor. Phys. Suppl. **191**, 192 (2011).

470 [40] L. Tubiana, E. Orlandini, and C. Micheletti, Phys. Rev. Lett. **107**, 188302 (2011).

471 [41] L. Tubiana, A. Rosa, F. Fragiacomo, and C. Micheletti, Macromolecules **46**, 3669 (2013).

472 [42] H. Flyvbjerg and H.G. Petersen, J. Chem. Phys. **91**, 461 (1989).

473 [43] M. Caraglio, F. Baldovin, B. Marcone, E. Orlandini, and A.L. Stella, ACS Macro Lett. **8**,
474 576 (2019).

475 [44] C. Micheletti and E. Orlandini, ACS Macro Lett. **3**, 876 (2014).

476 [45] R. Matthews, A.A. Louis, and J.M. Yeomans, EPL **89**, 20001 (2010).

477 [46] L. Huang and D.E. Makarov, J. Phys. Chem. A **111**, 10338 (2007).

[47] V. Narsimhan, C.B. Renner, and P.S. Doyle, ACS Macro Lett. **5**, 123 (2016).

Supplementary Information for "Diffusion of knots in nanochannel-confined DNA molecules"

Runfang Mao and Kevin D. Dorfman

Department of Chemical Engineering and Materials Science, University of Minnesota - Twin Cities, 421 Washington Ave. SE, Minneapolis, Minnesota 55455, United States

1. Supplementary Figures

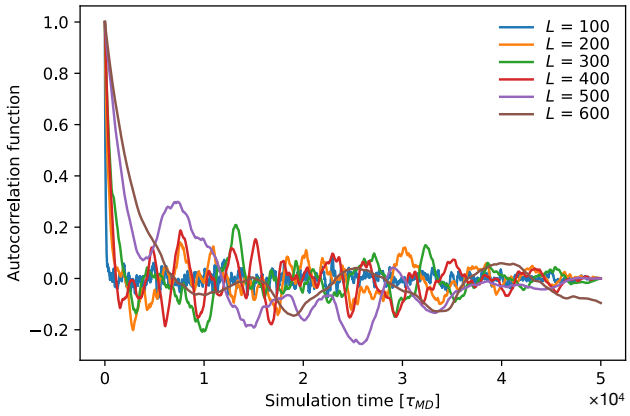

Figure S1. Autocorrelation function of radius of gyration for different chain length L. The autocorrelation time is 1×10^3, 1×10^3, 2×10^3, 5×10^3, 1×10^4, 2×10^4 τ_{MD} for L = 100, 200, 300, 400, 500 and 600, respectively.

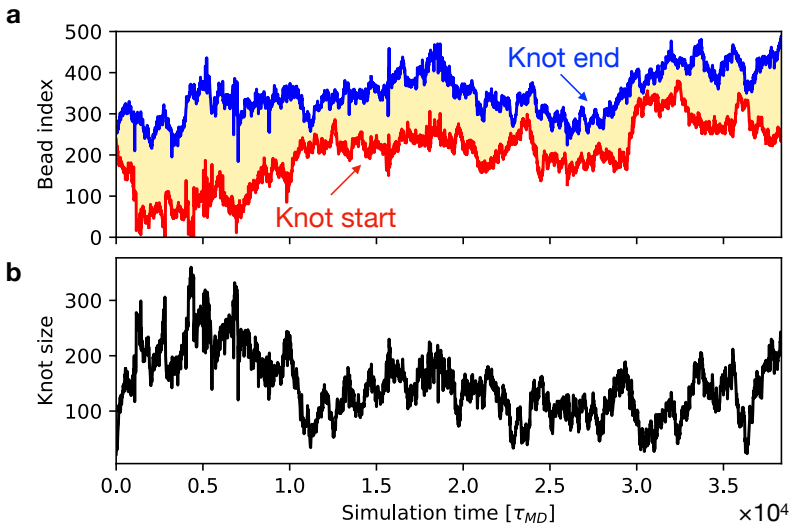

Figure S2. (a) The start (red line) and end (blue line) position of 3_1 trefoil knot along DNA chain in a MD simulation with different initial configuration at L = 500. (b) Knot size (i.e., the number of beads in the trefoil 3_1 knot core) as a function of simulation time calculated from upper panel.

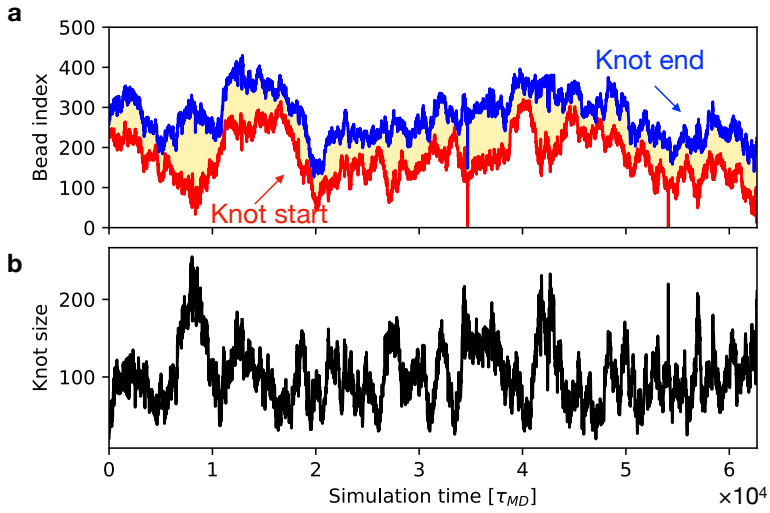

Figure S3. (a) The start (red line) and end (blue line) position of 3_1 trefoil knot along DNA chain in a MD simulation with different initial configuration at $L = 500$. (b) Knot size (i.e., the number of beads in the trefoil 3_1 knot core) as a function of simulation time calculated from upper panel.

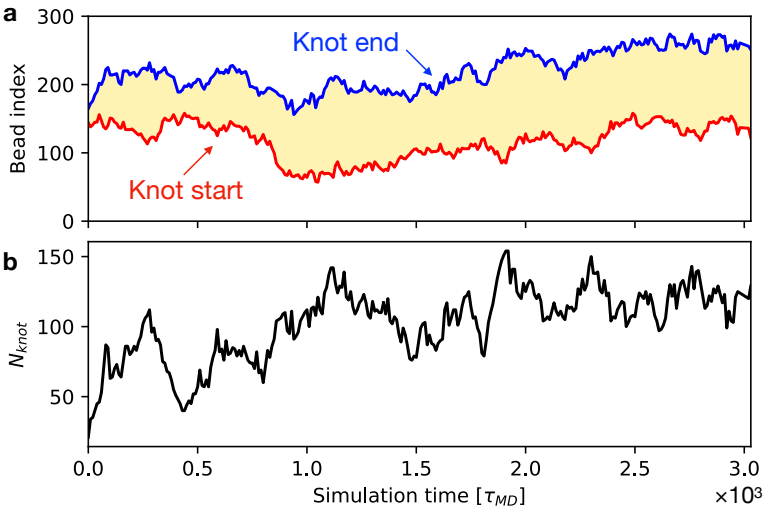

Figure S4. (a) The start (red line) and end (blue line) position of 3_1 trefoil knot along DNA chain in a MD simulation with total bead number $L = 300$. (b) The total bead number contained in the knot N_{knot} as a function of simulation time calculated from upper panel.

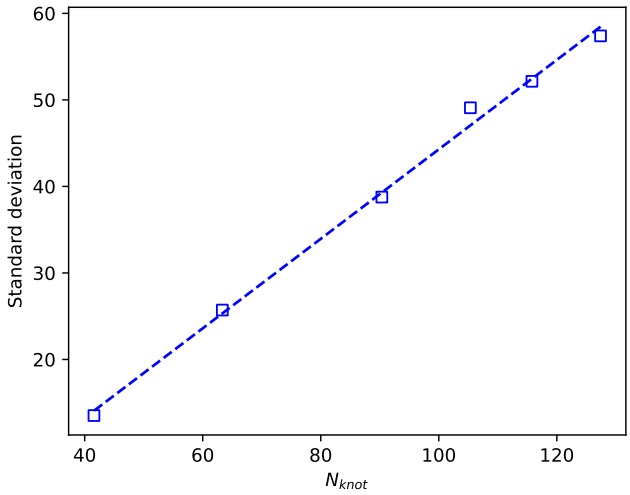

Figure S5. The standard deviation in N_{knot} as a function of N_{knot}.

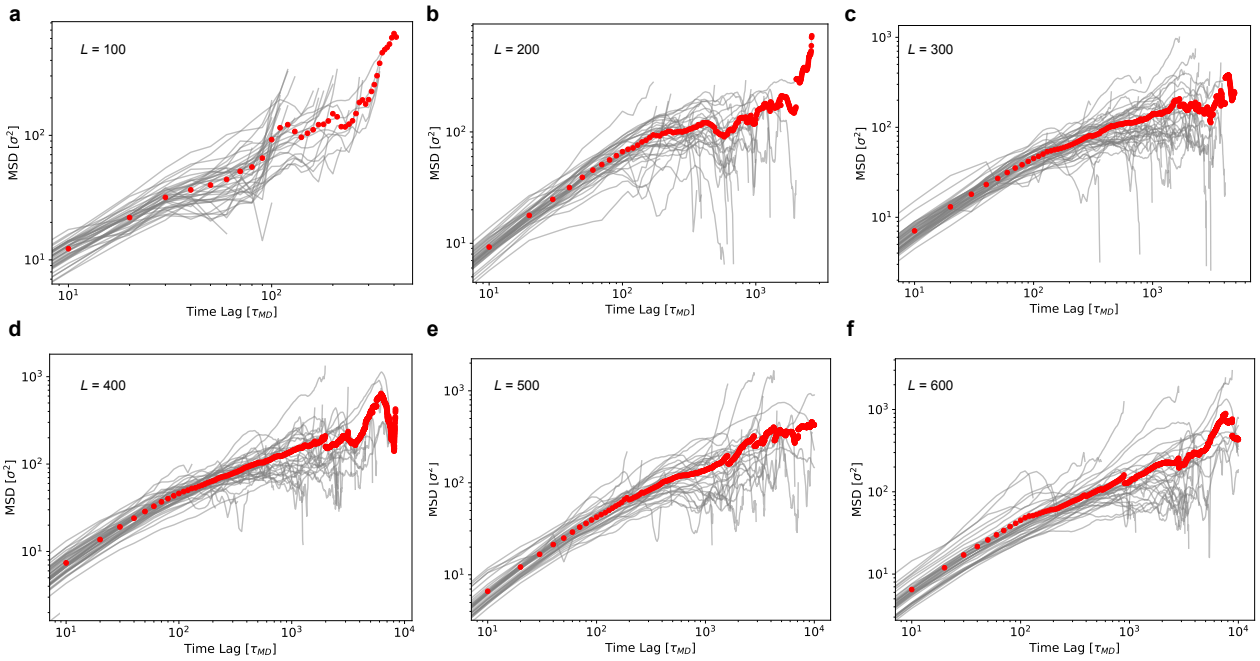

Figure S6. (a-f) The time-averaged MSDs (gray) for calculating ensemble-averaged MSD (red) from all simulation trajectories at L = 100, 200, 300, 400, 500, and 600, respectively.

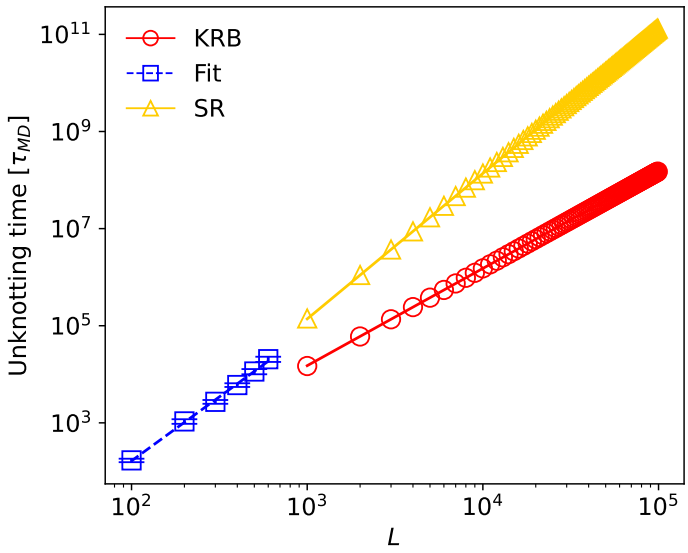

Figure S7. Unknotting time vs. L for varied chain length L = 1000 to 100000 with knot size Δ = 14. The predicted unknotting time for the knot region breathing (KRB) and the self-reptation (SR) is $\tau_{KRB} = 10^{-1.83}L^{2.0}$ and $\tau_{SR} = 10^{-3.85}L^{3.0}$.

2. Supplementary text

S1. Comparison of unknotting dynamics

As mentioned in the main text, we compared the unknotting dynamics with the reported knotting-unknotting dynamics of linear chain in nanochannel confinement.[1] In the compared reference paper, the chain contour length was reported as Lc = 3.6 μm, which corresponds to L = 300 with a bead diameter of σ = 12 nm in our simulation. Using the same assumption with Ref. 1, the characteristic simulation time $\tau_{MD} = \sigma(m/k_BT)^{1/2} = 6\pi\eta_{sol}\sigma^3/\varepsilon = 6\pi\eta_{sol}\sigma^3/(k_BT)$, where η_{sol} = 1 cP is the nominal water viscosity. At T = 300 K and σ = 12 nm, we obtained τ_{MD} = 8223 ns. Thus, the unknotting time in our simulation is about 3x10^3 τ_{MD} = 24 ms at L = 300 as shown in Figure 5. The reported value of the mean knot duration is around 5 ms for a similar channel size at D = 300 nm in Ref. 1. Considering the differences in the model and simulation setup, we can conclude that the unknotting time reported in our simulation is within the same order of magnitude compared to the results in Ref. 1.

References:

[1] Micheletti, C, and E Orlandini, ACS Macro Lett. **9**, 876 (2014).