

# Transfer-matrix calculations of DNA polymer micromechanics under tension and torque constraints

Artem K. Efremov,<sup>1,2,\*</sup> Ricksen S. Winardhi,<sup>1,3</sup> and Jie Yan<sup>1,2,3,†</sup>

<sup>1</sup>*Mechanobiology Institute, National University of Singapore, Singapore 117411*

<sup>2</sup>*Centre for Bioimaging Sciences, National University of Singapore, Singapore 117546*

<sup>3</sup>*Department of Physics, National University of Singapore, Singapore 117542*

(Dated: September 19, 2017)

Recent development of single-molecule manipulation technologies has made it possible to exert constant force and torque on individual DNA biopolymers to probe their elastic characteristics and structural stability. It has been previously shown that depending on the nature of applied mechanical constraints, DNA can exist in several forms including B-, L- and P-DNA. However, there is still a lack of understanding in how structural heterogeneity of DNA, which may naturally arise due to sequence-dependent DNA properties, protein binding or DNA damage, influence local stability of the above DNA states. To provide a more complete and detailed description of the DNA mechanics, we developed a theoretical framework based on transfer-matrix calculations and demonstrated how it can be used to predict the DNA behaviour upon application of a wide range of force and torque constraints. The resulting phase diagram shows DNA structural transitions that are in good agreement with previous experimental and theoretical studies. We further discuss how the constructed formalism can be extended to include local inhomogeneities in the DNA physical properties, thus making it possible to investigate the effect of DNA sequence as well as protein binding on DNA structural stability.

PACS numbers: 87.14.gk, 87.15.La, 87.15.A-, 82.37.Rs, 05.70.Fh

Keywords: DNA, phase transition, equilibrium thermodynamics, force-spectroscopy

## I. INTRODUCTION

Deoxyribonucleic acid (DNA) is a fundamental biomolecule carrying genetic information in all living organisms. It consists of two single-stranded polymers that are wound around each other forming a double helix structure. Both of these strands are made of a chain of basic building blocks, each containing a phosphate group, a deoxyribose sugar, and a nitrogenous base of one of the four distinct types A, T, G, C. Highly specific formation of hydrogen bonds between A:T and G:C base-pairs (bp) of the opposite DNA strands as well as the stacking interaction between aromatic rings of adjacent bases are the two major forces responsible for the stability of the double-helix structure of DNA.

In cells, DNA exists mainly in B-form – a right-handed duplex with a helical repeat  $h_B = 10.4$  bp [1]. Therefore, the number of times the two strands of a relaxed B-DNA of  $M$  base-pairs length twist around each other is:  $Lk_{0,B} = M/h_B$ . This quantity, usually referred to as the DNA linking number, is typically used to characterize the topological state of DNA that is often subjected to various mechanical constraints in living cells. Indeed, existing experimental data show that the conformation of the genomic DNA can be modulated by a large amount of different DNA-binding proteins, which either actively or passively change the DNA linking number. In particular, DNA-topoisomerases that can wind or unwind DNA

duplex cooperating with various DNA-architectural proteins, such as histones, are key elements responsible for the regulation of the linking number of a circular DNA or topologically isolated DNA domains *in vivo* [2].

In experiments using single-molecule manipulation techniques, it has been shown that if torque exerted on DNA surpasses a certain threshold value, the DNA frequently undergoes conformational changes to reduce the accumulated twist elastic energy either by collapsing into a supercoiled configuration or switching into an alternative structural form [3–9].

In the first scenario, DNA typically develops plectonemes – structures in which DNA is twisted into a helical braid-like conformation [10, 11]. In the second scenario, DNA experiences transition from B- to L- and P-DNA forms. L-DNA is a left-handed duplex with a helical repeat  $h_L$  of  $\sim 15$  bp, which occurs when DNA is placed under sufficiently large negative torsional stress ( $\lesssim -10$  pN·nm) [4, 8, 9]. Conversely, at high positive torques ( $\gtrsim 35$ -40 pN·nm) DNA switches into a right-handed P-DNA state that has a helical repeat  $h_P$  of  $\sim 3$  bp [4, 7, 12]. Although the exact molecular configurations of L- and P-DNA are not presently known, existing experimental studies suggest that L-DNA likely comprises a mixture of denatured and left-handed Z-DNA structures [4]. As for P-DNA, molecular dynamics simulations show that at high positive torques DNA undergoes complete denaturation that results in extruding bases [12].

Both the DNA conformation and interaction with DNA-binding proteins are determined by an intricate interplay between the elastic properties and structural stability of the DNA, mechanical constraints applied to it,

---

\* Corresponding author e-mail: [mbiay@nus.edu.sg](mailto:mbiay@nus.edu.sg)

† Corresponding author e-mail: [phyyj@nus.edu.sg](mailto:phyyj@nus.edu.sg)

and the nature of the nucleoprotein complexes that are formed. Detailed understanding of the contribution of each of these factors into the global DNA organization requires development of a systematic theoretical framework. While the previous theoretical studies have provided many important insights into the above problem [13–19], they were mainly based on the assumption of uniformity of the DNA physical properties without taking into account possible local variations, which may either result from the sequence-dependence of the DNA parameters or due to DNA segments' transitions into alternative structural states.

Here we present a new approach based on transfer-matrix calculations, which naturally takes into consideration local orientational and structural fluctuations of DNA. We demonstrate how this method can be used to provide a systematic description of the DNA behaviour under applied force / torque constraints, and discuss how local inhomogeneities in the DNA physical properties can be easily incorporated into it. Obtained results depicting the DNA behaviour under various mechanical constraints show good agreement with the existing theoretical and experimental studies.

In Sec. II, we outline the general theory and then apply it in Sec. III A to describe the force- and torque-dependent behaviour of B-DNA. In Sec. III B, the theory is further extended to characterize the structural stability of DNA and then used to estimate the phase boundaries between alternative DNA states. In Sec. III C, we demonstrate how the developed formalism can be utilized to predict behaviour of DNA with inhomogeneous physical properties. Finally, in Sec. IV, future applica-

tions of the theory to the understanding of the sequence-dependent DNA stability and DNA-protein interactions under force-torque constraints are discussed.

## II. GENERAL THEORY

To describe the behaviour of DNA under force and torque constraints, we modelled the DNA as a discretized polygonal chain consisting of short segments of equal length,  $b$  (Figure 1(a)). In order to accurately take into account local mechanical deformations, the value of  $b$  is chosen such that  $b \ll A$  and  $b \ll C$ , where  $A$  and  $C$  are the DNA bending and twisting persistence lengths. Each DNA segment is considered as a rigid body represented by a local Cartesian coordinate frame  $(\mathbf{x}_j, \mathbf{y}_j, \mathbf{z}_j)$ , which is generated by a three-dimensional rotation about the origin of a fixed lab coordinate system  $(\mathbf{x}_0, \mathbf{y}_0, \mathbf{z}_0)$ :  $\mathbf{x}_j = \mathbf{R}_j \mathbf{x}_0$ ,  $\mathbf{y}_j = \mathbf{R}_j \mathbf{y}_0$ , and  $\mathbf{z}_j = \mathbf{R}_j \mathbf{z}_0$ . Here  $j$  is the index of the corresponding DNA segment, and  $\mathbf{R}_j = \mathbf{R}_{\alpha_j} \mathbf{R}_{\beta_j} \mathbf{R}_{\gamma_j}$  denotes a rotation matrix resulting from the composition of three successive Euler revolutions through angles  $\alpha_j$ ,  $\beta_j$  and  $\gamma_j$  shown on Figure 1(b) (where  $\alpha_j, \gamma_j \in [0, 2\pi]$  and  $\beta_j \in [0, \pi]$ ). Since the Cartesian coordinate frame of each DNA segment  $(\mathbf{x}_j, \mathbf{y}_j, \mathbf{z}_j)$  is in a one-to-one relation to the respective rotational matrix  $\mathbf{R}_j$ , hereafter we simply use  $\mathbf{R}_j$  to designate the  $j^{\text{th}}$  DNA segment orientation.

In the general case, the expression for the total energy,  $E_{\text{tot}}$ , of DNA consisting of  $N$  segments includes several terms:

$$E_{\text{tot}}(\mathbf{R}_1, \dots, \mathbf{R}_N) = E_{\text{bend}} + E_{\text{twist}} + \Phi_f + \Phi_\tau \\ = \sum_{j=1}^{N-1} \left[ \frac{a}{2} (\mathbf{R}_j \mathbf{z}_0 - \mathbf{R}_{j+1} \mathbf{z}_0)^2 + \frac{c}{2} (2\pi \Delta T w_j(\mathbf{R}_j, \mathbf{R}_{j+1}))^2 \right] - bf \sum_{j=1}^N (\mathbf{z}_0 \cdot \mathbf{R}_j \mathbf{z}_0) - 2\pi\tau \Delta Lk \quad (1)$$

Here all of the energies are in  $k_B T$  units. The first sum in the above expression represents the bending,  $E_{\text{bend}}$ , and twisting,  $E_{\text{twist}}$ , elastic energies of the DNA resulting from the local contributions of adjacent DNA segments.  $a = A/b$  and  $c = C/b$  are dimensionless parameters describing the bending and twisting rigidities of DNA segments in the discretized model.  $\Delta T w_j(\mathbf{R}_j, \mathbf{R}_{j+1}) \approx \frac{1}{2\pi} \mathbf{R}_j \mathbf{z}_0 \cdot [\mathbf{R}_j \mathbf{x}_0 \times \mathbf{R}_{j+1} \mathbf{x}_0]$  is the local DNA twist.

The second sum corresponds to the potential energy,  $\Phi_f$ , arising from the force,  $f$ , applied to the DNA along  $\mathbf{z}_0$ -axis, while the last term takes into account the potential energy,  $\Phi_\tau$ , resulting from the torque,  $\tau$ , exerted on the DNA. In this formula, both  $f$  and  $\tau$  are scaled by  $k_B T$ ; therefore,  $f$  has a dimension of  $[1/L]$  and  $\tau$  is dimensionless.

$\Delta Lk$  is the DNA linking number change from the reference state, which is chosen to be a straight tor-

sionally relaxed B-DNA form. Using the Calugareanu-White's theorem [20, 21],  $\Delta Lk$  can be expressed as a sum of two components:  $\Delta Lk = \Delta T w + W r$ , where  $\Delta T w = \sum_{j=1}^{N-1} \Delta T w_j(\mathbf{R}_j, \mathbf{R}_{j+1})$  is the DNA total twist and  $W r$  is the DNA writhe number. In general,  $W r$  depends on the global DNA conformation and can be calculated using the Gauss double integral along the DNA contour (Appendix G, eq. (G1)). However, it is usually more convenient to use a simpler Fuller formula [22], which provides accurate estimations of the DNA writhe number for certain cases (see explanation following the equation):

$$W r^F = \frac{1}{2\pi} \int_0^L \frac{(\mathbf{z}_0 \cdot [\mathbf{t}(s) \times \dot{\mathbf{t}}(s)])}{1 + (\mathbf{t}(s) \cdot \mathbf{z}_0)} ds \quad (2)$$

Here  $L$  is the DNA contour length;  $\mathbf{t}(s)$  is the tangent vector of the DNA polymer parametrized by the arc

length  $s$ , and  $\dot{\mathbf{t}} = \frac{d\mathbf{t}}{ds}$ . It can be shown that  $Wr^F = Wr$  for those DNA conformations which can be obtained by a continuous deformation of DNA initially extended along  $\mathbf{z}_0$ -axis direction in such a way that the denominator of the integrand in eq. (2) never resets to zero for any of the intermediate DNA configurations [13, 22]. A nearly straight DNA, or DNA folded into a helical solenoid conformation or into a plectoneme aligned along  $\mathbf{z}_0$ -axis direction are examples satisfying this criterion [13, 16]. In other cases, however, the DNA writhe number  $Wr$  does not necessarily equal to  $Wr^F$ . In fact, it can be proved that  $Wr = Wr^F \bmod 2$  [16, 22, 23].

Despite the above uncertainty in evaluation of the writhe number for some DNA configurations [23, 24], it has been previously shown that the Fuller's formula can be used to recapitulate the experimentally observed DNA behaviour up to the onset of the torque-induced buckling transition when DNA starts to develop supercoiled plectoneme structures [14, 15]. Furthermore, the Fuller's formula has a great advantage since it allows the DNA writhe number to be calculated as a sum of local contributions from the DNA segments. This makes it possible to apply transfer-matrix calculations to predict the transition boundary between the extended and supercoiled DNA conformations and to evaluate changes in the values of the main observable parameters, such as the DNA extension or linking number, in response to the force and torque constraints.

Considering the above notes, we approximated the last term in eq. (1) describing the DNA torque-dependent potential energy,  $\Phi_\tau$ , as:

$$\Phi_\tau = -2\pi\tau\Delta Lk^F + \tau\lambda Wr^F \quad (3)$$

Here  $\Delta Lk^F = \Delta Tw + Wr^F$ , and the scaling parameter  $\lambda$

is introduced to account for the repulsive interaction energy between remote DNA sites brought into a close contact by the DNA buckling, which is expected to monotonically increase with the DNA writhe number. Since most of extended DNA conformations have writhe  $Wr^F \sim 0$ , the additional  $\tau\lambda Wr^F$  term makes a rather small contribution to the overall energy of such nearly-straight DNA states. Conversely, this additional term mainly affects the statistics of buckled DNA configurations appearing near the onset of the DNA buckling transition. As will be shown in Sec. III A and III B, under proper choice of the value of  $\lambda$ , this approximation results in a reasonable description of the supercoiling transition boundaries on the DNA phase diagram, which appears to be in good agreement with the existing experimental data.

By using the Fuller's formula (eq. (2)), it can be shown that the DNA linking number change,  $\Delta Lk^F$ , can be expressed as a sum of local DNA segments' contributions [13]:

$$\Delta Lk^F = \frac{1}{2\pi} \sum_{j=1}^{N-1} (\tilde{\alpha}_{j+1} + \tilde{\gamma}_{j+1} - \tilde{\alpha}_j - \tilde{\gamma}_j). \quad (4)$$

where  $\tilde{\alpha}_j$  and  $\tilde{\gamma}_j$  are the Euler angles of the  $j^{\text{th}}$  DNA segment from the extended range of  $(-\infty, +\infty)$ .

Substituting eq. (4) and (3) into eq. (1), it is then straightforward to see that eq. (1), in turn, transforms into a sum of local energy contributions of neighbouring DNA segments:

$$E_{\text{tot}}(\mathbf{R}_1, \dots, \mathbf{R}_N) = \sum_{j=1}^{N-1} E_j(\mathbf{R}_j, \mathbf{R}_{j+1}) - bf(\mathbf{z}_0 \cdot \mathbf{R}_N \mathbf{z}_0) \quad (5)$$

where the local energy terms  $E_j(\mathbf{R}_j, \mathbf{R}_{j+1})$  assume the following form:

$$E_j(\mathbf{R}_j, \mathbf{R}_{j+1}) = \frac{a}{2}(\mathbf{R}_j \mathbf{z}_0 - \mathbf{R}_{j+1} \mathbf{z}_0)^2 + \frac{c}{2}(2\pi\Delta Tw_j)^2 - bf(\mathbf{z}_0 \cdot \mathbf{R}_j \mathbf{z}_0) - \tau(2\pi - \lambda)\Delta Lk_j^F - \tau\lambda\Delta Tw_j \quad (6)$$

Here  $\Delta Tw_j = \Delta Tw_j(\mathbf{R}_j, \mathbf{R}_{j+1})$  and  $\Delta Lk_j^F = \frac{1}{2\pi}(\tilde{\alpha}_{j+1} + \tilde{\gamma}_{j+1} - \tilde{\alpha}_j - \tilde{\gamma}_j)$ .

Using eq. (5), the partition function of DNA,  $Z_{f,\tau}$ , under applied tension,  $f$ , and torque,  $\tau$ , can be found as:

$$\begin{aligned} Z_{f,\tau} &= \int d\mathbf{R}_1 \dots d\mathbf{R}_N e^{-E_{\text{tot}}(\mathbf{R}_1, \dots, \mathbf{R}_N)} \xi(\mathbf{R}_N, \mathbf{R}_1) = \\ &= \int d\mathbf{R}_1 \dots d\mathbf{R}_N \prod_{j=1}^{N-1} T(\mathbf{R}_j, \mathbf{R}_{j+1}) \times \sigma(\mathbf{R}_N, \mathbf{R}_1) \quad (7) \end{aligned}$$

where  $\xi(\mathbf{R}_N, \mathbf{R}_1)$  is a function that imposes specific boundary conditions on the orientations of the DNA ends;  $\sigma(\mathbf{R}_N, \mathbf{R}_1) = \xi(\mathbf{R}_N, \mathbf{R}_1) e^{bf(\mathbf{z}_0 \cdot \mathbf{R}_N \mathbf{z}_0)}$  and  $T(\mathbf{R}_j, \mathbf{R}_{j+1}) = e^{-E_j(\mathbf{R}_j, \mathbf{R}_{j+1})}$ . In the above formula the integration is carried out over all of the

possible DNA segments' orientations, i.e.,  $\int d\mathbf{R}_j = \int_0^{2\pi} d\alpha_j \int_0^{2\pi} d\gamma_j \int_0^\pi \sin\beta_j d\beta_j$ .

To simplify eq. (7), it is useful to recall that any square-integrable function,  $\psi$ , defined on SO(3) group of 3D rotation matrices parametrized by Euler angles  $\alpha, \beta, \gamma$  can be expanded into a series of orthogonal D-functions,  $D_{ml}^n(\alpha, \beta, \gamma)$  [25]. The expansion coefficients of this series can be found as  $\psi_{mln} = \frac{2n+1}{8\pi^2} \int d\mathbf{R} \psi(\mathbf{R}) \overline{D}_{ml}^n(\mathbf{R})$ , where for the sake of simplicity hereafter we use the following notations:  $\psi(\mathbf{R}) = \psi(\alpha, \beta, \gamma)$  and  $D_{ml}^n(\mathbf{R}) = D_{ml}^n(\alpha, \beta, \gamma)$ . The bar over function  $D_{ml}^n$  in the above formula denotes complex conjugation. Performing such an expansion for functions  $\sigma(\mathbf{R}_N, \mathbf{R}_1)$  and  $T(\mathbf{R}_j, \mathbf{R}_{j+1})$ , it can be shown that eq. (7) reduces to a mere summation of the respective expansion coefficients (see Appen-

lices A-C for more details):

$$Z_{f,\tau} = \sum_{\substack{m_1, \dots, m_N \\ l_1, \dots, l_N \\ n_1, \dots, n_N}} \left[ \prod_{j=1}^{N-1} T_{m_j, l_j, n_j}^{m_{j+1}, l_{j+1}, n_{j+1}} \times \sigma_{m_N, l_N, n_N}^{m_1, l_1, n_1} \right] \quad (8)$$

here

$$T_{m_j, l_j, n_j}^{m_{j+1}, l_{j+1}, n_{j+1}} = \frac{\sqrt{(2n_j+1)(2n_{j+1}+1)}}{8\pi^2} \int d\mathbf{R}_j d\mathbf{R}_{j+1} \bar{D}_{m_j, l_j}^{n_j}(\mathbf{R}_j) \times T(\mathbf{R}_j, \mathbf{R}_{j+1}) \times D_{m_{j+1}, l_{j+1}}^{n_{j+1}}(\mathbf{R}_{j+1}) \quad (9)$$

are the expansion coefficients of  $T(\mathbf{R}_j, \mathbf{R}_{j+1})$  function organized in a form of a multi-dimensional matrix,  $\mathbf{T} = [T_{m_j, l_j, n_j}^{m_{j+1}, l_{j+1}, n_{j+1}}]$ , and  $\boldsymbol{\sigma} = [\sigma_{m_N, l_N, n_N}^{m_1, l_1, n_1}]$  is the respective matrix of the expansion coefficients for the boundary condition function, which can be obtained using a similar formula.

It is possible to further streamline eq. (8) assuming that the boundary condition function  $\xi(\mathbf{R}_N, \mathbf{R}_1)$  pos-

sesses a symmetry about  $\mathbf{z}_0$ -axis of the lab coordinate system with respect to either the first or the last DNA segments' orientations – such that rotation of any of these two DNA segments about  $\mathbf{z}_0$ -axis does not change the value of the boundary condition function. For example, we may suppose that there are no restrictions on the possible orientations of the DNA end segments, in which case  $\xi(\mathbf{R}_N, \mathbf{R}_1) = 1$ . Alternatively, we may assume that the first and the last DNA segments are collinear to the  $\mathbf{z}_0$ -axis of the lab coordinate system, i.e.,  $\xi(\mathbf{R}_N, \mathbf{R}_1) = \delta(\mathbf{R}_1 \mathbf{z}_0 - \mathbf{z}_0) \delta(\mathbf{R}_N \mathbf{z}_0 - \mathbf{z}_0)$ , where  $\delta$  is Dirac delta function. Anyway, as soon as the boundary condition function has a symmetry about  $\mathbf{z}_0$ -axis, it can be shown that eq. (8) can be reduced to a simpler form, where the multi-dimensional transfer matrix  $\mathbf{T}$  is replaced by a conventional two-dimensional matrix  $\mathbf{S}$ , whose elements,  $S_{n_j n_{j+1}} = T_{0,0,n_j}^{0,0,n_{j+1}}$ , are defined by the following analytical expression (see Appendix B and the crucial relations in eq. (B9), (B14), (B19)-(B20) for more details):

$$S_{nn'}(a, b, c, \lambda) = \pi^2 \sqrt{(2n+1)(2n'+1)} e^{-a-c} \times \sum_{p,k,k',r} (2p+1)(2k+1)(2k'+1) e^{ir(\omega - \frac{\pi}{2})} I_r(\tau(1 - \frac{\lambda}{2\pi})) I_r(c\sqrt{1+\chi^2}) \mathcal{L}_r^p(-a) \mathcal{L}_r^k(-bf) \mathcal{L}_r^{k'}(0) \begin{pmatrix} p & k & n \\ -r & r & 0 \end{pmatrix}^2 \begin{pmatrix} p & k' & n' \\ -r & r & 0 \end{pmatrix}^2 \quad (10)$$

Here  $i$  is imaginary unit;  $\chi = \frac{\tau\lambda}{2\pi c}$  and  $\omega = \tan^{-1}(\chi)$ ;  $I_r$  is modified Bessel function of the first kind;  $\mathcal{L}_r^k(s) = \int_{-1}^1 d_{rr}^k(\cos^{-1} x) e^{-sx} dx$  is bilateral Laplace transform of the diagonal elements  $d_{rr}^k$  of small d-Wigner matrix;  $\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}$  are Wigner 3-j symbols.

Similarly, the multi-dimensional boundary condition matrix  $\boldsymbol{\sigma}$  becomes replaced by a two-dimensional matrix  $\mathbf{V}$ , whose elements are  $V_{nn'} = \sigma_{0,0,n}^{0,0,n'}$ . Using the expansion coefficients formula, it is easy to find that in the case of the first and last DNA segments  $\mathbf{z}_0$ -axis collinear boundary condition (see Appendix C, eq. (C12)-(C15) and eq. (D3)):

$$V_{nn'}(b) = \frac{1}{4\pi} e^{bf} \sqrt{(2n+1)(2n'+1)} \quad (11)$$

Alternatively, in the case of DNA ends free orientation boundary condition we have:

$$V_{nn'}(b) = 8\pi^2 \delta_{n'0} i_n(bf) \sqrt{2n+1} \quad (12)$$

Here  $i_n$  is modified spherical Bessel function of the first kind and  $\delta_{n'n}$  is the Kronecker delta ( $\delta_{n'n} = 1$  if  $n' = n$  and  $\delta_{n'n} = 0$ , otherwise).

Taking into account the above notes, it can be shown that the partition function for a homogeneous DNA polymer under applied force and torque constraints can be

computed as the trace of the product of the transfer,  $\mathbf{S}$ , and boundary condition,  $\mathbf{V}$ , matrices:

$$Z_{f,\tau} = \text{Tr}(\mathbf{S}^{N-1} \mathbf{V}) \quad (13)$$

Knowing the partition function  $Z_{f,\tau}$ , it is then straightforward to calculate the DNA force-extension curve under a constant torque ( $\tau = \tau_0$ ) and/or the torque-extension curve under a constant force ( $f = f_0$ ) as  $z_{\tau_0}(f) = \frac{\partial \ln(Z_{f,\tau=\tau_0})}{\partial f}$  and  $z_{f_0}(\tau) = \frac{\partial \ln(Z_{f,\tau})}{\partial \tau} |_{f=f_0}$ , respectively. The corresponding DNA linking number changes under the above mechanical constraints are  $\Delta Lk_{\tau_0}(f) = \frac{1}{2\pi} \frac{\partial \ln(Z_{f,\tau})}{\partial \tau} |_{\tau=\tau_0, \chi=\text{const}}$  and  $\Delta Lk_{f_0}(\tau) = \frac{1}{2\pi} \frac{\partial \ln(Z_{f=f_0,\tau})}{\partial \tau} |_{\chi=\text{const}}$  (here parameter  $\chi$  is treated as a constant). Since both the DNA extension and linking number change can be directly measured in single-molecule experiments, the above formulas provide a simple way to test the semi-flexible DNA model presented in this study.

We would like to note that while eq. (13) was derived assuming  $\mathbf{z}_0$ -axis symmetry of the boundary condition function, this expression can be also used to describe the behaviour of any sufficiently long DNA with arbitrary boundary conditions. To show this, matrices  $\mathbf{T}$  and  $\boldsymbol{\sigma}$

must be first re-written in two-dimensional forms:

$$T_{pp'} = T_{m,l,n}^{m',l',n'} \quad \text{and} \quad \sigma_{pp'} = \sigma_{m,l,n}^{m',l',n'} \quad (14)$$

Where the elements of two-dimensional matrices are enumerated by indexes  $p$  and  $p'$ , which relate to the old indexes as  $p(m, l, n) = m + l(2n + 1) + \frac{1}{3}n(4n^2 + 6n + 5)$  with a similar expression for  $p' = p'(m', l', n')$ . Then, substituting eq. (14) into eq. (8) and applying the notion of the matrix Jordan normal form, it can be demonstrated that in the case of a long DNA the partition function equals to  $Z_{f,\tau} \approx \kappa \nu_{max}^{N-1}$ , where  $\nu_{max}$  is the largest eigenvalue of the two-dimensional transfer matrix  $\mathbf{T}$  and  $\kappa$  is the proportionality constant. It is then clear from this expression for the DNA partition function that the boundary condition matrix  $\sigma$  only affects the value of the coefficient  $\kappa$ , which simply results in the DNA free energy offset by a fixed constant ( $-k_B T \ln \kappa$ ). Thus, behaviour of any sufficiently long DNA under force and torque constraints is independent from the exact form of the boundary condition function.

For example, the transfer-matrix calculations show that all of the quantitative results that will be discussed below differ by  $< 5\%$  if the DNA ends free orientation boundary condition (eq. (12)) is used instead of the  $\mathbf{z}_0$ -axis collinear boundary condition (eq. (11)) to estimate the partition function of  $\sim 5.1$  kbp long DNA. I.e., two completely opposite scenarios - partially fixed and completely free DNA ends' orientation, lead to virtually indistinguishable results. Therefore, it does not matter what kind of boundary condition is used in the transfer-matrix computations as soon as the contour length of DNA is much longer than its bending and twisting persistence lengths ( $L \gg A$  and  $L \gg C$ ).

Finally, it should be noted that in the case of  $\tau = 0$ , the transfer matrix elements defined by eq. (10) converge back to the worm-like chain model of a torsionally relaxed DNA, demonstrating consistency between the current work and previous theoretical studies (see Appendix D).

### III. RESULTS

#### A. The mechanical response of B-DNA

To probe how accurate the transfer-matrix approach is in recapitulating the transition boundary between extended and supercoiled DNA conformations, we first studied the mechanical response of B-DNA, assuming that it cannot switch into alternative structural states. In these calculations, the DNA bending and twisting persistence lengths were set equal to the experimentally measured values of  $A_B = 50$  nm and  $C_B = 95$  nm [4, 5, 26]. To correctly describe elastic deformations of DNA at small scales, the length of each DNA segment,  $b$ , was selected to be much shorter than the DNA bending and twisting persistence lengths:  $b = 0.5$  nm (i.e.,

each segment contains  $\sim 1.5$  base-pairs since each base-pair has rise of 0.33 nm in B-DNA form). Therefore, the dimensionless bending and twisting rigidities of the polygonal chain in the discretized model of DNA were  $a_B = A_B/b = 100$  and  $c_B = C_B/b = 190$ .

The value of the only unknown model parameter,  $\lambda$ , was adjusted to match the experimentally observed onset of the DNA buckling transition. It was found that by setting  $\lambda$  from 4.1 to 4.4, the model can accurately depict the experimentally measured values of the applied force and torque at which DNA begins to collapse into a supercoiled configuration (see Figure 2).  $\lambda_B = 4.3$  was used for B-DNA for the rest of the calculations performed in this study. As for the DNA size, we set it equal to  $\sim 5.1$  kbp (totally 3400 DNA segments) in all our computations.

Substituting the above model parameters into eq. (10), (11) and (13), we plotted the B-DNA force-extension curves,  $z_{\tau_0}(f)$ , at various torque constraints ( $\tau = \tau_0$ ) as shown in Figure 3(a). As can be seen from the graph, formation of supercoiled DNA structures (manifested by a steep DNA extension decrease) begins to take place only when the applied torque exceeds the threshold value of  $\sim 5$  pN·nm. In contrast, at small torques of  $0 < \tau < 5$  pN·nm the DNA extension is only slightly shorter compared to that of a torsionally relaxed DNA ( $\tau = 0$ ) due to the chiral bending fluctuations previously discussed by Nelson and Marko [11, 15, 27].

While Figure 3(a) demonstrates B-DNA behaviour at positive torques, it should be noted that DNA folding into supercoiled conformations occurs both at positive and negative torques in a symmetric manner, which is expected since in these calculations B-DNA is not allowed to switch into alternative structural forms (i.e., L- or P-DNA). This symmetry can be clearly seen from Figure 3(b), showing that in addition to a rapid extension drop, the development of supercoiled DNA domains is also accompanied by a steep change in the DNA linking number,  $\Delta Lk_{\tau_0}(f)$ . The DNA torque-extension ( $z_{f_0}(\tau)$ ) and linking number change curves ( $\Delta Lk_{f_0}(\tau)$ ) under various force constraints ( $f = f_0$ ) exhibit a very similar symmetric behaviour with respect to the applied torque, see Figures 3(c-d).

The above data indicate that the approximation of the torque-dependent potential energy term,  $\Phi_\tau$ , by eq. (3) yields reasonable results, reminding those observed in single-DNA manipulation experiments, see Figure 2 in ref. [3].

While the transfer-matrix calculations correctly predict position of the DNA supercoiling transition boundaries on the DNA phase diagram (see Figures 2 and 6), it should be noted that due to the use of the Fuller's approximation supercoiled DNA structures formed after the DNA buckling transition in the presented model do not necessarily correspond to DNA plectonemes observed in real experiments. To gain further insights into the shape of these DNA structures, we carried out Metropolis-Monte-Carlo simulations based on eq. (1)-(6), see Appendix G for more details. It was found that although

the DNA configurations generated by the Metropolis-Monte-Carlo algorithm (Figure 4(a)) assumed compact conformations, they lack the expected supercoiled DNA plectonemes.

Moreover, DNA writhe number estimations obtained for these DNA configurations using the double Gauss integral and the Fuller's formula show a drastic difference as indicated in Figure 4(a). We further examined the contributions of different DNA segments into the total DNA writhe number by applying the both formula to one of the generated DNA configurations (Figure 4(b), please refer to Appendix G for the details of the DNA segments' writhe number calculations). It can be seen from the figure that the two calculations give similar results for DNA segments that are nearly aligned along the force direction. In contrast, DNA segments that are aligned in the opposite direction have an expected near-zero contribution to the DNA writhe number calculated using the double Gauss integral, while the Fuller's formula leads to erroneously large values due to the vanishing denominator in eq. (2). Such overestimation of the DNA writhe number by the Fuller's formula leads to a strong decrease of the DNA total energy (eq. (1)) at sufficiently high torques, which results in preference of incorrectly collapsed DNA conformations over other DNA configurations.

These results suggest that the torque-driven collapse of the DNA extension (Figure 3(a-c)) and large linking number change (Figure 3(b-d)) under low force are likely caused by the formation of these incorrect conformations. Therefore, while the Fuller's formula can be used to accurately predict the behaviour of DNA in extended configurations, it does not seem to correctly describe supercoiled DNA conformations. Hence, the application of the presented theoretical framework, and any calculations using the Fuller's formula in general, should be restricted up to the onset of the DNA buckling transition.

In the next subsection, we further extend the transfer-matrix approach to investigate the structural stability of DNA by taking into consideration that it can switch between several alternative forms, such as B-, L- or P-DNA, which have very distinct physical properties [4, 7–9, 12].

## B. The structural stability of DNA

DNA transition from B- to L- or P-DNA states is determined by the base-pairing energy difference between the respective DNA structures, which can be written as  $\mu_u - 2\pi\tau\Delta lk_0^{(u)}$ . Here  $\mu_u$  is the energy difference at zero torque, corresponding to the transition from B-DNA to a particular state denoted by index  $u = L, P$  or  $B$ ; and  $\Delta lk_0^{(u)} = lk_{0,u} - lk_{0,B}$ , where  $lk_{0,u} = \pm h_u^{-1}$  is the relaxed linking number of the respective DNA form per single base-pair and  $h_u$  is the helical repeat of DNA in this form. In the above formula, the sign of  $lk_{0,u}$  is positive for right-handed DNA helical structures (like B- and P-

DNA) and negative for left-handed structures (L-DNA). Since we use B-DNA as the reference point it is clear that  $\mu_B = 0$  and  $\Delta lk_0^{(B)} = 0$ .

Taking into account the aforementioned notes, it can be shown that the transfer matrix,  $\mathbf{S}_u$ , for DNA segments in state  $u$  can be obtained via multiplication of eq. (10) by an exponential factor containing the base-pairing energy difference (see Appendix E for details):

$$(\mathbf{S}_u)_{nn'} = e^{-q(\mu_u - 2\pi\tau\Delta lk_0^{(u)})} \mathbf{S}_{nn'}(a_u, b_u, c_u, \lambda_u) \quad (15)$$

Here  $q$  is the number of base-pairs in individual DNA segments;  $b_u = q\rho_u$  is the length of the DNA segments in state  $u$ , and  $\rho_u$  is the base-pair rise in the corresponding DNA form.  $a_u = A_u/b_u$ ,  $c_u = C_u/b_u$  and  $\lambda_u$  are the model parameters describing the respective physical properties of DNA segments in states  $u = B, L$  or  $P$ .

Using eq. (15), the DNA partition function can be found as (see Appendix E):

$$Z_{f,\tau} = \text{Tr}(\mathbf{U}\hat{\mathbf{S}}^{N-1}\hat{\mathbf{V}}) \quad (16)$$

Where  $\mathbf{U} = (\mathbf{I} \ \mathbf{I} \ \mathbf{I})$  is a block-matrix comprising three identity matrices,  $\mathbf{I}$ . As for the transfer matrix,  $\hat{\mathbf{S}}$ , and boundary condition matrix,  $\hat{\mathbf{V}}$ , they have the following forms:

$$\hat{\mathbf{S}} = \begin{pmatrix} \mathbf{S}_B & \mathbf{S}_B & \mathbf{S}_B \\ \mathbf{S}_L & \mathbf{S}_L & \mathbf{S}_L \\ \mathbf{S}_P & \mathbf{S}_P & \mathbf{S}_P \end{pmatrix}, \quad \hat{\mathbf{V}} = \begin{pmatrix} \mathbf{V}_B \\ e^{-q(\mu_L - 2\pi\tau\Delta lk_0^{(L)})} \mathbf{V}_L \\ e^{-q(\mu_P - 2\pi\tau\Delta lk_0^{(P)})} \mathbf{V}_P \end{pmatrix} \quad (17)$$

Here blocks  $\mathbf{S}_u$  are defined by eq. (15), and blocks  $\mathbf{V}_u = \mathbf{V}(b_u)$  are described by previously introduced eq. (11) and (12) depending on the DNA ends boundary conditions.

In all our calculations, the size of the DNA segments was set to be equal to  $\sim 1.5$  bp for all of the DNA structures and the DNA length was  $\sim 5.1$  kbp (a total of 3400 segments in the discretized polymer chain representing DNA). For the rest of the L- and P-DNA parameters we used the values listed below.

**L-DNA.** Compared to B-DNA, L-DNA has much smaller bending and twisting persistence lengths of  $A_L = 7$  nm and  $C_L = 15$  nm [8, 9]. Taking into account that the contour length of L-DNA per base-pair is  $\sim 1.35$  times larger than that of B-DNA ( $b_L = 1.35 \times b_B = 0.675$  nm) [8, 19], the dimensionless bending and twisting rigidities of DNA segments in the discretized polymer chain model equal to  $a_L = A_L/b_L = 10.4$  and  $c_L = C_L/b_L = 22.2$ .

As for the base-pairing energy difference  $\mu_L$ , it is believed to be in a range of a few  $k_B T$  [17–19]. A fixed value  $\mu_L = 5.0 k_B T$  was used in the calculations presented below. Furthermore, based on the structural data for B- and L-DNA, it can be estimated that  $\Delta lk_0^{(L)} = -h_L^{-1} - h_B^{-1} = -0.16$ . As will be seen from the computation results, with these parameters our model can correctly describe the DNA transition from B- to L-form,

which takes place at the experimentally measured torque values of  $-11$  to  $-10$  pN·nm [4, 8, 9].

Finally, due to the lack of information in the existing literature regarding the transition boundary between extended and supercoiled L-DNA conformations, the value of parameter  $\lambda_L$  was simply set equal to the value of  $\lambda_B$  for B-DNA :  $\lambda_L = \lambda_B = 4.3$ . We note that the choice of  $\lambda_L$  only affects the supercoiling transition boundary of L-DNA but does not affect the B-to-L-DNA transition.

**P-DNA.** The bending and twisting persistence lengths of P-DNA have been previously estimated to be  $A_P = 15$  nm and  $C_P = 25$  nm [19]. Since the contour length of P-DNA per base-pair is  $\sim 1.7$  times larger than that of B-DNA ( $b_P = 1.7 \times b_B = 0.85$  nm), we have  $a_P = A_P/b_P = 17.6$  and  $c_P = C_P/b_P = 29.4$  for the dimensionless bending and twisting rigidities of P-DNA segments.

The relaxed linking number difference between the B- and P-DNA structures per base-pair is  $\Delta lk_0^{(P)} = h_P^{-1} - h_B^{-1} = 0.24$ . As for the transition energy difference  $\mu_P$ , it was estimated based on comparison of the calculated phase boundary between B- and P-DNA states with the experimentally measured one, which takes place at  $\sim 35$ - $40$  pN·nm torque [4, 7]. It was found that the latter value can be reproduced in our computations by setting  $\mu_P = 17.8 k_B T$ , in good agreement with the value of  $15 k_B T$  estimated in an earlier theoretical work [19].

Lastly, the existing experimental data show that the onset of the P-DNA transition from the extended into a supercoiled conformation takes place at  $\sim 20$  pN forces and  $35 - 50$  pN·nm torques [12]. Using the transfer-matrix calculations, we found that a value of  $\lambda_P = -0.5$  results in the P-DNA supercoiling transition in the same range of applied forces and torques.

Substituting the above model parameters into eq. (15)-(17), we calculated the DNA extension and linking number change as functions of applied force and torque constraints. Figure 5(a) shows the force-extension curves,  $z_{\tau_0}(f)$  (top), and force-linking number change curves,  $\Delta Lk_{\tau_0}(f)$  (bottom), plotted at different negative (left) and positive (right) torques  $\tau_0$ . As can be seen from the figure, for sufficiently small torques ( $\tau_0 \in [-5, 5]$  pN·nm) the calculated curves are virtually similar to those obtained for a torsionally relaxed DNA ( $\tau_0 = 0$  pN·nm) – the DNA extension demonstrates practically the same response to the applied force, and only a slight linking number change is observed with the increasing absolute value of the torque.

At larger negative torques of  $\tau_0 \in [-10, -5]$  pN·nm, the DNA behaviour still reminds that of a torsionally relaxed DNA at high forces. However, when the applied force drops below a certain threshold value the DNA extension experiences a steep collapse accompanied by simultaneous decrease in the DNA linking number, resembling the typical transition of negatively wound B-DNA into a supercoiled conformation observed in single-DNA manipulation experiments.

Application of even stronger negative torques ( $\tau_0 < -10$  pN·nm) results in an abrupt change of the  $z_{\tau_0}(f)$  and  $\Delta Lk_{\tau_0}(f)$  curves' pattern indicated by  $\sim 1.3$  times longer DNA extension at high forces and a simultaneous large DNA linking number decrease (by  $\sim 900$ ). These results are consistent with the DNA transition from the right-handed B-form into a longer left-handed L-DNA structure at  $\sim -11$  pN·nm torque observed in experimental studies [4, 8, 9]. Similar to B-DNA, L-DNA switches from the extended to a supercoiled conformation as soon as the applied force drops below a certain threshold, which is slightly larger for L-DNA as compared to the B-DNA case due to its higher elasticity.

As for the DNA behavior at positive torques, it practically mirrors the one described above for negative torques with the only difference being that B-DNA switches into alternative P-DNA state at a much higher torque of  $\sim 35$ - $40$  pN·nm than in the case of B-to-L-DNA transition. From Figure 5(a), right panels, it can be seen that B-to-P-DNA transformation is accompanied by a large elongation of the DNA contour length with a simultaneous abrupt increase in the DNA linking number by  $\sim 1400$ . The estimated B-to-P-DNA transition boundary at  $\sim 35$ - $40$  pN·nm is consistent with previous experimental observations [4, 7].

At larger torques ( $> 35$ - $40$  pN·nm) where P-DNA structure is stable, our calculations predict that supercoiling of P-DNA occurs when the applied force drops below  $\sim 20$  pN, resulting in further DNA linking number increase. Both the value of the B-P base-pairing energy difference ( $\mu_P = 17.8 k_B T$ ) and the transition boundary of P-DNA from the extended to a supercoiling conformation estimated by our transfer-matrix method are similar to those obtained by Marko and Neukirch using a different theoretical approach [19].

Figure 5(b) shows the torque-extension curves,  $z_{f_0}(\tau)$ , and the torque-linking number change curves,  $\Delta Lk_{f_0}(\tau)$ , under various force constraints ( $f = f_0$ ). It can be seen that in the case of  $f_0 < 0.5$  pN, all  $z_{f_0}(\tau)$  curves have symmetric profiles with respect to both positive and negative torques. However, as soon as the applied force increases above  $f_0 \sim 0.5$ - $0.7$  pN, this symmetry breaks due to B-DNA switching into L-DNA state at sufficiently large negative torques. Further increase of the applied force beyond  $f_0 \sim 15$  pN results in appearance of B-to-P transition, which takes place at  $\sim 35$ - $40$  pN·nm torque.

Using the above theoretical data, we plotted the DNA phase diagram showing the transition boundaries between different DNA states (see Figure 6). The supercoiling transition boundary for each DNA structure was assumed to pass through the points on the force-torque diagram where the DNA extension experiences  $\sim 50\%$  drop with respect to the value predicted by the worm-like chain model for the corresponding DNA structure. As for the boundaries between alternative DNA forms (B, L or P), they were defined as the set of points ( $f_0, \tau_0$ ) at which  $\sim 50\%$  of the DNA segments are in L- or P-DNA states, respectively. For this purpose, the total number of DNA

segments,  $N_u$ , in state  $u = \text{L or P}$  at various values of the applied force ( $f = f_0$ ) and torque ( $\tau = \tau_0$ ) was calculated as  $(N_u)_{f_0, \tau_0} = -\frac{1}{q} \frac{\partial \ln(Z_{f, \tau})}{\partial \mu_u} \Big|_{f=f_0, \tau=\tau_0}$ .

As can be seen from Figure 6, the resulting phase diagram demonstrates good agreement with the existing experimental data [4, 5, 7–9, 12, 28] as well as with the previously reported theoretical studies [17–19].

Here we emphasize again that although the transfer-matrix calculations can predict the onset of the B-, L- and P-DNA supercoiling transition boundaries, the conformations of the collapsed DNA structures do not necessarily correspond to experimentally observed DNA plectonemes due to the use of the Fuller’s formula approximation.

### C. DNA with structural heterogeneity

In the above sections, it was shown how the transfer-matrix method can be applied to describe the behaviour of DNA with homogeneous physical properties under force and torque constraints. It is clear that the constructed formalism can be also used to take into consideration any mechanical and structural heterogeneity along DNA. Indeed, according to eq. (16) the DNA partition function is determined by the product of transfer matrices defined for individual DNA segments. Therefore, any local deformation or non-uniformity of physical properties of DNA changes only the elements of the transfer matrix corresponding to the respective DNA segment. It is then straightforward to extend the developed formalism to the general case through site-dependent transfer matrices,  $\hat{\mathbf{S}}_j$ , by replacing the matrices product  $\hat{\mathbf{S}}^{N-1}$  with  $\prod_{j=1}^{N-1} \hat{\mathbf{S}}_j$  in eq. (16).

To demonstrate how such approach can be utilized to predict the behaviour of DNA with local inhomogeneity, we performed a number of calculations for 5.1 kbp DNA that has a small insert of 24 bp, or 51 bp, or 102 bp in the middle part. This insert is characterized by a lower energy cost of the B-to-L transition ( $\mu_L = 2.0 k_B T$ ) comparing to that of the rest of the DNA ( $\mu_L = 5.0 k_B T$ ). This way, a local structural inhomogeneity was introduced into the DNA. The values of the other model parameters,  $a_u$ ,  $b_u$ ,  $c_u$  and  $\lambda_u$  (where  $u = \text{B, L or P}$ ), were kept unchanged in the computations.

Figure 7 shows the results of the transfer-matrix calculations, demonstrating how the linking number of a mechanically stretched DNA having the above mentioned local structural inhomogeneity changes with the applied torque. As can be seen from the figure, in the absence of the insert, the B-to-L transition of the DNA takes place at a negative torque of  $\sim -11$  pN·nm; whereas, in the case of the structural inhomogeneity, the DNA experiences a sequence of two B-to-L transitions - the first one at a small torque, which is then followed by the second at  $\sim -11$  pN·nm. The first transition corresponds to the insert switching from B- to L-DNA form, while the sec-

ond transition indicates the B-to-L transformation of the rest of the DNA.

Interestingly, very similar torque-linking number change curves were reported in a recent single-molecule study [9], where the behaviour of 4.6 kbp DNA molecules containing various small inserts in the middle part prone to L-DNA formation was studied at different force and torque constraints. Such similarity between our model predictions and the reported experimental data suggests that the transfer-matrix formalism constructed in our study correctly describes the effect of local DNA inhomogeneities. This makes it possible to employ in future the transfer-matrix calculations to extract the values of the parameters describing the physical properties of different DNA structures based on the results of single-DNA manipulation experiments.

## IV. DISCUSSION

In summary, we have developed a new method based on the transfer-matrix technique to evaluate the conformational changes of DNA subjected to force and torque constraints.

Although the transfer-matrix calculations has been previously used in theoretical studies of DNA, those theories either did not explicitly consider local DNA deformations [17] or were mainly focused on the behavior of a torsionally relaxed DNA [29–31], accurately describing the DNA behaviour only at high forces or at zero torque. In contrast, the major advantage of our method is that it explicitly takes into account local orientational and structural fluctuations of DNA, thereby allowing us to predict the DNA behaviour in a wide range of applied mechanical constraints. In spite of the need to introduce a new model parameter (i.e.,  $\lambda$  parameter in eq. (3)), this advantage makes it a unique and useful approach that provides description of the DNA mechanical response with minimal modifications to the polymer model of DNA.

It should be noted that  $\lambda$  parameter has a simple physical interpretation. Metropolis-Monte-Carlo simulations show that while the Fuller’s formula correctly estimates the DNA writhe number of extended DNA configurations, it provides wrong results for supercoiled DNA conformations, leading to a wrong statistical sampling of the latter in the partition function integral (eq. (7)). This in turn leads to disbalance between the extended and supercoiled DNA configurations, resulting in the DNA buckling transition taking place at forces higher than those measured in single-DNA manipulation studies. Thus, on the one hand, introduction of  $\lambda$  parameter into the model is intended to correct the disbalance between extended and supercoiled DNA conformations caused by the usage of the Fuller’s formula. On the other hand, the energy term corresponding to  $\lambda$  parameter also accounts for the contribution of the electrostatic repulsive interaction energy between remote DNA sites brought into a close contact by the DNA writhing, which has a profound effect



on the DNA supercoiling transition [19].

Here we again would like to stress that while the parameter  $\lambda$  is used in the model to shift the transition boundary between extended and supercoiled DNA conformations to the experimentally measured position on the phase diagram, it does not affect the behaviour of extended or nearly-extended DNA because the writhe number of such DNA configurations is typically very small. Furthermore,  $\lambda$  parameter neither affects the phase boundaries between B-, L- and P-DNA structures, as these transitions are mainly governed by the parameters  $\mu_u$  and  $\Delta l k_0^{(u)}$  ( $u = B, L$  or  $P$ ). Thus,  $\lambda$  can be thought of as a parameter responsible solely for the positioning of the DNA supercoiling transition boundaries on the DNA phase diagram.

A recent work by Marko and Neukirch [19] adopts a different strategy to tackle this problem by modeling DNA as a homogeneous polymer comprised of extended and supercoiled regions, which are treated as topologically separated but mechanically coupled domains. The simplicity of such approach has made it possible to obtain many important insights into the structural stability of DNA at various mechanical constraints. However, the requirement of the DNA homogeneity makes it difficult to apply this type of model to situations where DNA in-

homogeneities naturally arise as a result of the sequence dependent properties of DNA or due to binding of proteins. Our transfer-matrix method fills this gap since it is built based on local DNA deformations, without the need to separate the DNA into different topological domains. Because of these advantages, the formalism developed in our study may provide a wide scope of potential applications, such as investigation of sequence-dependent DNA mechanical response and exploration of interactions between DNA and architectural proteins [9, 32–38]. As such, the theory developed in this work can be conveniently applied to cases involving excited or permanent DNA heterogeneities caused by sequence-dependent variations of the DNA physical properties or local structural transitions of DNA as well as site-specific protein binding or DNA damage.

## V. ACKNOWLEDGEMENTS

This research was funded by Singapore Ministry of Education Academic Research Fund Tier 3 (MOE2012-T3-1-001), Tier 2 (MOE2013-T2-1-154) and the National Research Foundation through the Mechanobiology Institute Singapore to JY.

- 
- [1] J. C. Wang, *Proc. Natl. Acad. Sci. USA* **76**, 200 (1979).
  - [2] M. S. Luijsterburg, M. F. White, R. van Driel, and R. T. Dame, *Crit. Rev. Biochem. Mol. Biol.* **43**, 393 (2008).
  - [3] T. R. Strick, J.-F. Allemand, D. Bensimon, A. Bensimon, and V. Croquette, *Science* **271**, 1835 (1996).
  - [4] Z. Bryant, M. D. Stone, J. Gore, S. B. Smith, N. R. Cozzarelli, and C. Bustamante, *Nature* **424**, 338 (2003).
  - [5] S. Forth, C. Deufel, M. Y. Sheinin, B. Daniels, J. P. Sethna, and M. D. Wang, *Phys. Rev. Lett.* **100**, 148301 (2008).
  - [6] F. Mosconi, J. F. Allemand, D. Bensimon, and V. Croquette, *Phys. Rev. Lett.* **102**, 078301 (2009).
  - [7] M. Y. Sheinin and M. D. Wang, *Phys. Chem. Chem. Phys.* **11**, 4800 (2009).
  - [8] M. Y. Sheinin, S. Forth, J. F. Marko, and M. D. Wang, *Phys. Rev. Lett.* **107**, 108102 (2011).
  - [9] F. C. Oberstrass, L. E. Fernandes, and Z. Bryant, *Proc. Natl. Acad. Sci. USA* **109**, 6106 (2012).
  - [10] A. V. Vologodskii and N. R. Cozzarelli, *Annu. Rev. Biophys. Biomol. Struct.* **23**, 609 (1994).
  - [11] J. F. Marko, *Physica A* **418**, 126 (2015).
  - [12] J. F. Allemand, D. Bensimon, R. Lavery, and V. Croquette, *Proc. Natl. Acad. Sci. USA* **95**, 14152 (1998).
  - [13] B. Fain, J. Rudnick, and S. Ostlund, *Phys. Rev. E* **55**, 7364 (1997).
  - [14] C. Bouchiat and M. Mezard, *Phys. Rev. Lett.* **80**, 1556 (1998).
  - [15] J. D. Moroz and P. Nelson, *Macromolecules* **31**, 6333 (1998).
  - [16] C. Bouchiat and M. Mezard, *arXiv* (2000).
  - [17] A. Sarkar, J.-F. Léger, D. Chatenay, and J. F. Marko, *Phys. Rev. E* **63**, 051903 (2001).
  - [18] J. F. Marko, *Phys. Rev. E* **76**, 021926 (2007).
  - [19] J. F. Marko and S. Neukirch, *Phys. Rev. E* **88**, 062722 (2013).
  - [20] G. Calugareanu, *Czech. Math. J.* **11**, 588 (1961).
  - [21] J. H. White, *Am. J. Math.* **91**, 693 (1969).
  - [22] F. B. Fuller, *Proc. Natl. Acad. Sci. USA* **75**, 3557 (1978).
  - [23] V. Rossetto and A. Maggs, *arXiv* (2003).
  - [24] S. Neukirch and E. L. Starostin, *Phys. Rev. E* **78**, 041912 (2008).
  - [25] I. Gel'fand, R. Minlos, and Z. Shapiro, *Representations of the rotation and Lorentz groups and their applications*. (Pergamon Press, New York, 1963).
  - [26] C. Bustamante, J. F. Marko, E. D. Siggia, and S. Smith, *Science* **265**, 1599 (1994).
  - [27] J. D. Moroz and P. Nelson, *Proc. Natl. Acad. Sci. USA* **94**, 14418 (1997).
  - [28] C. Deufel, S. Forth, C. R. Simmons, S. Dejjosha, and M. D. Wang, *Nat. Methods* **4**, 223 (2007).
  - [29] J. Yan and J. F. Marko, *Phys. Rev. E* **68**, 011905 (2003).
  - [30] J. Yan, R. Kawamura, and J. F. Marko, *Phys. Rev. E* **71**, 061905 (2005).
  - [31] P. Ranjith, P. B. S. Kumar, and G. I. Menon, *Phys. Rev. Lett.* **94**, 138102 (2005).
  - [32] X. Yue, C. Hu, Q. Yu-Jie, A. K. Efremov, L. Ming, O. Zhong-Can, L. Dong-Sheng, and Y. Jie, *Chin. Phys. B* **23**, 068702 (2014).
  - [33] H. You, X. Zeng, Y. Xu, C. J. Lim, A. K. Efremov, A. T. Phan, and J. Yan, *Nucleic Acids Res.* **42**, 8789 (2014).
  - [34] A. K. Efremov, Y. Qu, H. Maruyama, C. J. Lim, K. Takeyasu, and J. Yan, *J. Biol. Chem.* **290**, 15770 (2015).

- [35] R. S. Winardhi, W. Fu, S. Castang, Y. Li, S. L. Dove, and J. Yan, *Nucleic Acids Res.* **40**, 8942 (2012).
- [36] R. S. Winardhi, R. Gulvady, J. L. Mellies, and J. Yan, *J. Biol. Chem.* **289**, 13739 (2014).
- [37] R. S. Winardhi, J. Yan, and L. J. Kenney, *Biophys. J.* **109**, 1321 (2015).
- [38] S. Le, H. Chen, P. Cong, J. Lin, P. Droge, and J. Yan, *Sci. Rep.* **3**, 3508 (2013).
- [39] G. B. Arfken and H. J. Weber, *Mathematical methods for physicists*. (Elsevier Academic Press, Cambridge, 2005).
- [40] H. Zhang and J. F. Marko, *Phys. Rev. E* **82**, 051906 (2010).
- [41] J. F. Marko and E. D. Siggia, *Macromolecules* **28**, 8759 (1995).
- [42] A. Messiah, *Quantum mechanics, volume 2*. (North-Holland Publishing Company, Amsterdam, 1965).
- [43] W. Rudin, *Functional analysis*. (McGraw-Hill, Inc., Singapore, 1991).
- [44] P. R. Halmos, *Measure theory*. (Springer-Verlag, New York, 1974).
- [45] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).
- [46] M. Levitt, *J. Mol. Biol.* **170**, 723 (1983).
- [47] K. Klenin and J. Langowski, *Biopolymers* **54**, 307 (2000).
- [48] A. V. Vologodskii and J. F. Marko, *Biophys. J.* **73**, 123 (1997).

## Appendix A: Orthogonal D-functions

To calculate the partition function of DNA, we use a semiflexible polymer chain model in which DNA is represented by a discretized polygonal chain consisting of short segments of equal length,  $b$ . The global DNA conformation is then uniquely determined by the orientations of all of the DNA segments, which can be described in terms of Euler rotation matrices,  $(\mathbf{R}_1, \dots, \mathbf{R}_N)$ , where  $N$  is the total number of segments in the polygonal chain. Here each matrix  $\mathbf{R}_j$  shows the orientation of the local Cartesian coordinate system,  $(\mathbf{x}_j, \mathbf{y}_j, \mathbf{z}_j)$ , attached to the  $j^{\text{th}}$  DNA segment relative to the fixed lab coordinate frame  $(\mathbf{x}_0, \mathbf{y}_0, \mathbf{z}_0)$ . Thus, collection  $(\mathbf{R}_1, \dots, \mathbf{R}_N)$  of all of the DNA segments rotation matrices completely specifies the global DNA conformation.

As was shown in the main text, using the Fuller's approximation for the DNA writhe number, the total energy of DNA can be represented as a sum of local energies  $E_j(\mathbf{R}_j, \mathbf{R}_{j+1})$  associated with the vertices connecting neighbouring DNA segments. In this case, calculation of the DNA partition function can be greatly simplified by reducing the numerous repetitive integrations of the product of  $e^{-E_j(\mathbf{R}_j, \mathbf{R}_{j+1})}$  functions to a mere multiplication of the respective transfer matrices, see eq. (8), (13) and (16) in the main text. To prove these formulas and to find out the expression for the transfer matrix elements (eq. (10) and (15)), we will use several famous results from the group theory.

First of all, it should be noted that every Euler matrix,  $\mathbf{R}$ , corresponds to a unique element from  $SO(3)$  group of all rotations about the origin of three-dimensional Euclidean space. It is convenient to parametrize  $SO(3)$  group by the three Euler angles  $\alpha$ ,  $\beta$  and  $\gamma$  (see Figure 1(b)), in which case every rotation matrix,  $\mathbf{R}$ , is determined by the respective set  $(\alpha, \beta, \gamma)$  of Euler angles ( $\alpha, \gamma \in [0, 2\pi]$  and  $\beta \in [0, \pi]$ ).

Second, from the group theory it is known that any square-integrable function defined on  $SO(3)$  group can be expanded into a series of orthogonal functions  $D_{m,l}^n$ , which have the following canonical form, see p. 101 in [25]:

$$D_{m,l}^n(\alpha, \beta, \gamma) = e^{-im\alpha} P_{m,l}^n(\cos \beta) e^{-il\gamma} \quad (\text{A1})$$

Where  $n, m, l$  are integers such that  $n \geq 0$  and  $-n \leq m, l \leq n$ ;  $\alpha, \beta$  and  $\gamma$  are Euler angles, parametrizing  $SO(3)$  group; and  $P_{m,l}^n$  are polynomials, which relate to the elements of so-called small Wigner d-matrix,  $d_{m,l}^n$ , as  $P_{m,l}^n(\cos \beta) = i^{m-l} d_{m,l}^n(\beta)$ .

Functions  $D_{m,l}^n$  and polynomials  $P_{m,l}^n$  possess a number of important properties, which we will extensively use below to derive the formulas for the transfer matrix elements and DNA partition function.

First, by substituting  $(\alpha, \beta, \gamma) = (0, 0, 0)$  into eq. (A1) and taking into account that  $d_{m,l}^n(0) = \delta_{ml}$ , we get:

$$D_{m,l}^n(0, 0, 0) = P_{m,l}^n(1) = i^{m-l} d_{m,l}^n(0) = \delta_{ml} \quad (\text{A2})$$

Here  $\delta_{ml}$  is the Kronecker delta ( $\delta_{ml} = 1$  if  $m = l$  and  $\delta_{ml} = 0$ , otherwise).

Furthermore, since for any indexes  $n \geq 0$  and  $-n \leq m, l \leq n$ :  $d_{m,l}^n(\beta)$  are real functions obeying the following symmetric relations  $d_{m,l}^n(\beta) = (-1)^{l-m} d_{l,m}^n(\beta) = d_{-l,-m}^n(\beta)$ , it is not very hard to see that:

$$(-1)^{l-m} \overline{P_{m,l}^n}(x) = P_{m,l}^n(x) = P_{l,m}^n(x) = P_{-m,-l}^n(x) \quad (\text{A3})$$

Where the bar over the function denotes the complex conjugate.

Combining together eq. (A1) and (A3), we obtain:

$$\overline{D_{m,l}^n}(\mathbf{R}) = (-1)^{m-l} D_{-m,-l}^n(\mathbf{R}) \quad (\text{A4})$$

Here and below for the sake of formulas simplicity we use  $D_{m,l}^n(\mathbf{R})$  notation to address functions  $D_{m,l}^n(\alpha, \beta, \gamma)$ , where  $\mathbf{R}$  is the Euler rotation matrix corresponding to angles  $(\alpha, \beta, \gamma)$ .

Using eq. (A1), (A3) and (A4), it is straightforward to show that:

$$D_{m,l}^n(\mathbf{R}^{-1}) = \overline{D}_{l,m}^n(\mathbf{R}) \quad (\text{A5})$$

Here matrix  $\mathbf{R}^{-1}$  corresponding to Euler angles  $(\pi - \gamma, \beta, \pi - \alpha)$  is the inverse of matrix  $\mathbf{R}$  (i.e.,  $\mathbf{R}^{-1}\mathbf{R} = \mathbf{R}\mathbf{R}^{-1} = \mathbf{I}$ , where  $\mathbf{I}$  is the  $3 \times 3$  identity matrix:  $I_{ml} = \delta_{ml}$ ).

Next, functions  $D_{m,l}^n$  obey the following important multiplication rules [25]:

$$D_{m,l}^n(\mathbf{R}_1\mathbf{R}_2) = \sum_{k=-n}^n D_{m,k}^n(\mathbf{R}_1) D_{k,l}^n(\mathbf{R}_2) \quad (\text{A6})$$

and

$$D_{m_1,l_1}^{n_1}(\mathbf{R}) D_{m_2,l_2}^{n_2}(\mathbf{R}) = \sum_n \langle n_1 n_2 m_1 m_2 | n(m_1 + m_2) \rangle \langle n_1 n_2 l_1 l_2 | n(l_1 + l_2) \rangle D_{m_1+m_2, l_1+l_2}^n(\mathbf{R}) \quad (\text{A7})$$

Where  $\langle n_1 n_2 m_1 m_2 | n_3 m_3 \rangle$  are Clebsh-Gordan coefficients. For the sake of the formulas simplicity and compactness, below we will use Winger 3-j symbols instead of Clebsh-Gordan coefficients, which relate to each other as:

$$\begin{pmatrix} n_1 & n_2 & n_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = \frac{(-1)^{n_1 - n_2 - m_3}}{\sqrt{2n_3 + 1}} \langle n_1 n_2 m_1 m_2 | n_3(-m_3) \rangle \quad (\text{A8})$$

The final important property of  $D_{m,l}^n$  functions required for the DNA transfer matrix derivation is their orthogonality, which was mentioned in the beginning of this appendix. Namely, it can be shown that [25]:

$$\int d\mathbf{R} \overline{D}_{m_1, l_1}^{n_1}(\mathbf{R}) D_{m_2, l_2}^{n_2}(\mathbf{R}) = \frac{8\pi^2}{2n_1 + 1} \delta_{n_1 n_2} \delta_{m_1 m_2} \delta_{l_1 l_2} \quad (\text{A9})$$

Where the integration in the above formula is carried out over all of the possible combinations of the Euler angles  $(\alpha, \beta, \gamma)$ :

$$\int d\mathbf{R} = \int_0^{2\pi} d\alpha \int_0^{2\pi} d\gamma \int_0^\pi \sin \beta d\beta \quad (\text{A10})$$

Orthogonality and completeness of  $D_{m,l}^n$  functions makes it possible to use them as a Hilbert basis in the space of square-integrable functions,  $F(\alpha, \beta, \gamma) = F(\mathbf{R})$ , defined on SO(3) group [25]. Therefore, any such function,  $F(\mathbf{R})$ , can be expanded into the following series:

$$F(\mathbf{R}) = \sum_{n=0}^{\infty} \sum_{m, l=-n}^n F_n^{ml} D_{m,l}^n(\mathbf{R}) \quad (\text{A11})$$

Where the expansion coefficients  $F_n^{ml}$  are:

$$F_n^{ml} = \frac{2n+1}{8\pi^2} \int d\mathbf{R} \overline{D}_{m,l}^n(\mathbf{R}) F(\mathbf{R}) \quad (\text{A12})$$

Analogously, for any square-integrable function  $F(\mathbf{R}, \mathbf{R}')$ , where  $\mathbf{R}$  and  $\mathbf{R}'$  are two rotation matrices, we have:

$$F(\mathbf{R}, \mathbf{R}') = \sum_{n, n'=0}^{\infty} \sum_{m, l=-n}^n \sum_{m', l'=-n'}^{n'} F_{nn'}^{mm' ll'} D_{m,l}^n(\mathbf{R}) \overline{D}_{m', l'}^{n'}(\mathbf{R}') \quad (\text{A13})$$

Where the expansion coefficients  $F_{nn'}^{mm' ll'}$  are:

$$F_{nn'}^{mm' ll'} = \frac{(2n+1)(2n'+1)}{(8\pi^2)^2} \int d\mathbf{R} d\mathbf{R}' \overline{D}_{m,l}^n(\mathbf{R}) F(\mathbf{R}, \mathbf{R}') D_{m', l'}^{n'}(\mathbf{R}') \quad (\text{A14})$$

With all of the above formulas at hand, it is now straightforward to deduct the expressions for the DNA partition function and the transfer matrix elements discussed in the main text.

## Appendix B: Elements of the DNA transfer matrix

Let's now apply eq. (A11)-(A14) to obtain expansion formulas for a couple of functions that will be later used to derive the transfer matrix for the DNA partition function calculations.

First of all, we will find the expansion series of the exponential function containing the potential energy terms corresponding to the DNA mechanical stretching and twisting under the applied load,  $f$ , and torque,  $\tau$ :  $e^{bf(\mathbf{z}_0 \cdot \mathbf{R}\mathbf{z}_0) + \tau(2\pi - \lambda)\Delta Lk^F(\mathbf{R}, \mathbf{R}')}$ . Here  $\mathbf{R}$  and  $\mathbf{R}'$  are two Euler matrices describing the orientations of neighbour DNA segments (DNA segment corresponding to matrix  $\mathbf{R}$  is followed by one corresponding to matrix  $\mathbf{R}'$ ); and  $\Delta Lk^F(\mathbf{R}, \mathbf{R}') = \frac{1}{2\pi}(\tilde{\alpha}' + \tilde{\gamma}' - \tilde{\alpha} - \tilde{\gamma})$  is the local contribution of the two DNA segments to the total change in the DNA linking number based on the Fuller's formula approximation, see more details in ref. [13, 14]. In the latter expression,  $\tilde{\alpha}, \tilde{\gamma}, \tilde{\alpha}', \tilde{\gamma}'$  are the Euler angles from the extended range of  $(-\infty, \infty)$  corresponding to matrices  $\mathbf{R}$  and  $\mathbf{R}'$ , which relate to usual Euler angles  $(\alpha, \beta, \gamma)$  and  $(\alpha', \beta', \gamma')$  as:

$$\alpha = \tilde{\alpha} \bmod 2\pi, \quad \gamma = \tilde{\gamma} \bmod 2\pi, \quad \alpha' = \tilde{\alpha}' \bmod 2\pi, \quad \gamma' = \tilde{\gamma}' \bmod 2\pi \quad (\text{B1})$$

Since  $\beta$  is the angle between  $\mathbf{z}_0$ -axis of the lab coordinate system and  $\mathbf{z}$ -axis of the system generated by Euler rotations  $(\alpha, \beta, \gamma)$  it is obvious that  $(\mathbf{z}_0 \cdot \mathbf{R}\mathbf{z}_0) = (\mathbf{z}_0 \cdot \mathbf{z}) = \cos \beta$ , see Figure 1(b). Furthermore, since the size of the DNA segments,  $b$ , is chosen to be much smaller than the bending,  $A$ , and twisting,  $C$ , persistence lengths of DNA ( $b \ll A$  and  $b \ll C$ ) it is clear that the coordinate frame corresponding to matrix  $\mathbf{R}'$  is only slightly rotated relative to the coordinate frame corresponding to matrix  $\mathbf{R}$ . In this case:

$$2\pi\Delta Lk^F(\mathbf{R}, \mathbf{R}') = \tilde{\alpha}' + \tilde{\gamma}' - \tilde{\alpha} - \tilde{\gamma} \approx \sin(\alpha' + \gamma' - \alpha - \gamma) \quad (\text{B2})$$

Taking into account the above notes, it is easy to see that:

$$e^{bf(\mathbf{z}_0 \cdot \mathbf{R}\mathbf{z}_0) + \tau(2\pi - \lambda)\Delta Lk^F(\mathbf{R}, \mathbf{R}')} = e^{bf \cos \beta + \tau(1 - \frac{\lambda}{2\pi}) \sin(\alpha' + \gamma' - \alpha - \gamma)} \quad (\text{B3})$$

To find the expansion formula for the above function, it is convenient to use Jacobi-Anger equation (p. 687, [39]):

$$e^{iq \cos \varphi} = \sum_{n=-\infty}^{+\infty} i^n J_n(q) e^{in\varphi} \quad (\text{B4})$$

Where  $J_n(x)$  are Bessel functions of the first kind;  $i$  is imaginary unit and  $q$  is an arbitrary constant. Substituting  $\psi = \frac{\pi}{2} - \varphi$  and  $\rho = iq$  into eq. (B4), we get:

$$e^{\rho \sin \psi} = \sum_{n=-\infty}^{+\infty} i^{-n} I_n(\rho) e^{in\psi} \quad (\text{B5})$$

Where  $I_n(x) = i^{-n} J_n(ix)$  are modified Bessel functions of the first kind, which have the following properties:  $I_{-n}(x) = I_n(x)$  and  $I_n(-x) = (-1)^n I_n(x)$ , see p. 714 in ref. [39]. By using functions  $I_n(x)$ , it also is possible to re-write eq. (B4) in another more convenient form:

$$e^{\rho \cos \varphi} = \sum_{n=-\infty}^{+\infty} I_n(\rho) e^{in\varphi} \quad (\text{B6})$$

From eq. (A14), (B3) and (B5) we have:

$$\begin{aligned} F_{nn'}^{mm'l'l'} &= \frac{(2n+1)(2n'+1)}{(8\pi^2)^2} \int d\mathbf{R}d\mathbf{R}' \bar{D}_{m,l}^n(\mathbf{R}) e^{bf(\mathbf{z}_0 \cdot \mathbf{R}\mathbf{z}_0) + \tau(2\pi - \lambda)\Delta Lk^F(\mathbf{R}, \mathbf{R}')} D_{m',l'}^{n'}(\mathbf{R}') = \frac{(2n+1)(2n'+1)}{(8\pi^2)^2} \times \\ &\times \int d\mathbf{R}d\mathbf{R}' \bar{P}_{m,l}^n(\cos \beta) e^{bf \cos \beta} \times P_{m',l'}^{n'}(\cos \beta') \times \sum_{k=-\infty}^{+\infty} i^{-k} I_k(\tau(1 - \frac{\lambda}{2\pi})) e^{i(m-k)\alpha + i(l-k)\gamma - i(m'-k)\alpha' - i(l'-k)\gamma'} = \\ &= \delta_{mm'} \delta_{ll'} \delta_{ml} \times \frac{1}{4} (2n+1)(2n'+1) i^{-m} I_m(\tau(1 - \frac{\lambda}{2\pi})) \mathcal{L}_m^n(-bf) \mathcal{L}_m^{n'}(0) \end{aligned} \quad (\text{B7})$$

Here functions  $\mathcal{L}_m^n(s)$  are bilateral Laplace transforms of  $P_{m,m}^n$  polynomials / elements of small Wigner d-matrix:

$$\mathcal{L}_m^n(s) = \int_{-1}^1 P_{m,m}^n(x) e^{-sx} dx = \int_{-1}^1 d_{m,m}^n(\cos^{-1} x) e^{-sx} dx \quad (\text{B8})$$

Substituting eq. (B7) into eq. (A13), we finally get the desired expansion formula:

$$e^{bf(\mathbf{z}_0 \cdot \mathbf{R}\mathbf{z}_0) + \tau(2\pi - \lambda)\Delta Lk^F(\mathbf{R}, \mathbf{R}')} = \frac{1}{4} \sum_{n, n', m} (2n+1)(2n'+1) i^{-m} I_m\left(\tau \left(1 - \frac{\lambda}{2\pi}\right)\right) \mathcal{L}_m^n(-bf) \mathcal{L}_m^{n'}(0) D_{m, m}^n(\mathbf{R}) \overline{D}_{m, m}^{n'}(\mathbf{R}') \quad (\text{B9})$$

The second function,  $e^{-\frac{\alpha}{2}(\mathbf{R}\mathbf{z}_0 - \mathbf{R}'\mathbf{z}_0)^2 - \frac{\alpha}{2}(2\pi\Delta Tw(\mathbf{R}, \mathbf{R}'))^2 + \tau\lambda\Delta Tw(\mathbf{R}, \mathbf{R}')}$ , which we want to expand, contains the remaining energy terms from eq. (6) in the main manuscript text corresponding to the local elastic deformation of DNA taking place between neighbour DNA segments, whose orientations in space are described by Euler matrices  $\mathbf{R}$  and  $\mathbf{R}'$ . Here  $\Delta Tw(\mathbf{R}, \mathbf{R}')$  is the local contribution of the two DNA segments to the DNA total twist number (i.e.,  $\Delta Tw(\mathbf{R}, \mathbf{R}')$  is the twist angle between the coordinate systems corresponding to Euler matrices  $\mathbf{R}$  and  $\mathbf{R}'$  normalized to  $2\pi$  radians).

Before proceeding to the expansion series derivation it should be first noted that the above function depends only on the product  $\mathbf{R}^{-1}\mathbf{R}'$  of Euler matrices. Indeed, from the definition of rotation matrices it follows that  $(\mathbf{R}\mathbf{z}_0 - \mathbf{R}'\mathbf{z}_0)^2 = 2 - 2(\mathbf{R}\mathbf{z}_0 \cdot \mathbf{R}'\mathbf{z}_0) = 2 - 2(\mathbf{z}_0 \cdot \mathbf{R}^{-1}\mathbf{R}'\mathbf{z}_0)$  and  $\Delta Tw(\mathbf{R}, \mathbf{R}') = \Delta Tw(\mathbf{I}, \mathbf{R}^{-1}\mathbf{R}')$ , where  $\mathbf{I}$  is the identity rotation matrix corresponding to the lab coordinate system ( $\alpha_0, \beta_0, \gamma_0 = 0$ ). In other words, the twisting angle between the coordinate systems corresponding to Euler matrices  $\mathbf{R}$  and  $\mathbf{R}'$  as well as the bending angle between their  $\mathbf{z}$ -axes depend only on the relative orientation of the two coordinate systems in space and is independent from their exact alignment with respect to the lab coordinate frame ( $\mathbf{x}_0, \mathbf{y}_0, \mathbf{z}_0$ ). Thus, the expansion series of the above exponential function can be found in two steps. First, we will consider the special case in which the coordinate system corresponding to matrix  $\mathbf{R}$  is identical to the lab coordinate system ( $\mathbf{R} = \mathbf{I}$ ), and the coordinate frame corresponding to matrix  $\mathbf{R}'$  is only slightly rotated relative to it. Second, by substituting  $\mathbf{R}' \rightarrow \mathbf{R}^{-1}\mathbf{R}'$  into the formula obtained for the special case and using eq. (A6), we will get the desired expansion series for the general case.

Let again  $(\alpha', \beta', \gamma')$  be the Euler angles corresponding to matrix  $\mathbf{R}'$ . Then taking into account the above notes, for the special case of  $\mathbf{R} = \mathbf{I}$  we have:  $(\mathbf{R}\mathbf{z}_0 - \mathbf{R}'\mathbf{z}_0)^2 = 2 - 2(\mathbf{z}_0 \cdot \mathbf{R}'\mathbf{z}_0) = 2 - 2\cos\beta'$  and  $\Delta Tw(\mathbf{R}, \mathbf{R}') = \Delta Tw(\mathbf{I}, \mathbf{R}')$ . Furthermore, since the coordinate frame corresponding to matrix  $\mathbf{R}'$  is only slightly rotated relative to the lab coordinate system because  $b \ll A$  and  $b \ll C$ , it is clear that:

$$2\pi\Delta Tw(\mathbf{I}, \mathbf{R}') \approx \alpha' + \gamma' \approx \sin(\alpha' + \gamma') \quad \text{and} \quad (2\pi\Delta Tw(\mathbf{I}, \mathbf{R}'))^2 \approx 2 - 2\cos(\alpha' + \gamma') \quad (\text{B10})$$

Thus, in the special case of  $\mathbf{R} = \mathbf{I}$ :

$$e^{-\frac{\alpha}{2}(\mathbf{R}\mathbf{z}_0 - \mathbf{R}'\mathbf{z}_0)^2 - \frac{\alpha}{2}(2\pi\Delta Tw(\mathbf{R}, \mathbf{R}'))^2 + \tau\lambda\Delta Tw(\mathbf{R}, \mathbf{R}')} = e^{-\frac{\alpha}{2}(\mathbf{z}_0 - \mathbf{R}'\mathbf{z}_0)^2 - \frac{\alpha}{2}(2\pi\Delta Tw(\mathbf{I}, \mathbf{R}'))^2 + \tau\lambda\Delta Tw(\mathbf{I}, \mathbf{R}')} = e^{-a-c} e^{a\cos\beta' + c\cos(\alpha' + \gamma') + \frac{\tau\lambda}{2\pi}\sin(\alpha' + \gamma')} \quad (\text{B11})$$

Applying eq. (A12) and (B6) to the above expression, it is rather straightforward to show that:

$$\begin{aligned} F_{n', l'}^{m', l'} &= \frac{2n'+1}{8\pi^2} \int d\mathbf{R}' \overline{D}_{m', l'}^{n'}(\mathbf{R}') e^{-\frac{\alpha}{2}(\mathbf{z}_0 - \mathbf{R}'\mathbf{z}_0)^2 - \frac{\alpha}{2}(2\pi\Delta Tw(\mathbf{I}, \mathbf{R}'))^2 + \tau\lambda\Delta Tw(\mathbf{I}, \mathbf{R}')} = \\ &= \frac{2n'+1}{8\pi^2} e^{-a-c} \int d\mathbf{R}' \overline{P}_{m', l'}^{n'}(\cos\beta') e^{a\cos\beta'} \times \sum_{k=-\infty}^{+\infty} I_k\left(c\sqrt{1+\chi^2}\right) e^{i(m'+k)\alpha' + i(l'+k)\gamma' - ik\omega} = \\ &= \delta_{m', l'} \times \frac{1}{2} (2n'+1) e^{-a-c} e^{im'\omega} I_{m'}\left(c\sqrt{1+\chi^2}\right) \mathcal{L}_{m'}^{n'}(-a) \end{aligned} \quad (\text{B12})$$

Where  $\chi = \frac{\tau\lambda}{2\pi c}$  and  $\omega = \tan^{-1}(\chi)$ .

Substituting eq. (B12) into eq. (A11), we obtain the desired expansion formula for the special case of  $\mathbf{R} = \mathbf{I}$ :

$$e^{-\frac{\alpha}{2}(\mathbf{z}_0 - \mathbf{R}'\mathbf{z}_0)^2 - \frac{\alpha}{2}(2\pi\Delta Tw(\mathbf{I}, \mathbf{R}'))^2 + \tau\lambda\Delta Tw(\mathbf{I}, \mathbf{R}')} = \frac{1}{2} e^{-a-c} \sum_{n', m'} (2n'+1) e^{im'\omega} I_{m'}\left(c\sqrt{1+\chi^2}\right) \mathcal{L}_{m'}^{n'}(-a) D_{m', m'}^{n'}(\mathbf{R}') \quad (\text{B13})$$

To extend this formula to the general case, we simply need to put  $\mathbf{R}' \rightarrow \mathbf{R}^{-1}\mathbf{R}'$  and use the previously mentioned

multiplication property of  $D_{m,l}^n$  functions (eq. (A6)). By doing so, we get:

$$e^{-\frac{a}{2}(\mathbf{R}\mathbf{z}_0 - \mathbf{R}'\mathbf{z}_0)^2 - \frac{c}{2}(2\pi\Delta Tw(\mathbf{R}, \mathbf{R}'))^2 + \tau\lambda\Delta Tw(\mathbf{R}, \mathbf{R}')} = \frac{1}{2}e^{-a-c} \sum_{n,m,l} (2n+1) e^{-im\omega} I_m\left(c\sqrt{1+\chi^2}\right) \mathcal{L}_m^n(-a) D_{l,m}^n(\mathbf{R}) \overline{D}_{l,m}^n(\mathbf{R}') \quad (\text{B14})$$

Here, in addition to the indexes change  $n' \rightarrow n$  and  $m' \rightarrow -m$ , we also used eq. (A4) and (A5) to re-write the expansion formula in the form of eq. (A13).

Having the expansion series for the exponential functions containing both the global and local energy terms, it is now simple enough to obtain the general formula for the DNA transfer matrix. Namely, multiplying eq. (B9) by eq. (B14), and using twice the multiplication rule for  $D_{m,l}^n$  functions (eq. (A7)) together with eq. (A8), we have:

$$\begin{aligned} e^{-E(\mathbf{R}, \mathbf{R}')} &= e^{-\frac{a}{2}(\mathbf{R}\mathbf{z}_0 - \mathbf{R}'\mathbf{z}_0)^2 - \frac{c}{2}(2\pi\Delta Tw(\mathbf{R}, \mathbf{R}'))^2 + bf(\mathbf{z}_0 \cdot \mathbf{R}\mathbf{z}_0) + \tau(2\pi - \lambda)\Delta Lk^F(\mathbf{R}, \mathbf{R}') + \tau\lambda\Delta Tw(\mathbf{R}, \mathbf{R}')} = \\ &= \frac{1}{8}e^{-a-c} \sum_{n,m,l,k,k',r} (2n+1)(2k+1)(2k'+1) i^{-r} e^{-im\omega} I_m\left(c\sqrt{1+\chi^2}\right) I_r\left(\tau\left(1 - \frac{\lambda}{2\pi}\right)\right) \times \\ &\quad \times \mathcal{L}_m^n(-a) \mathcal{L}_r^k(-bf) \mathcal{L}_r^{k'}(0) \times D_{l,m}^n(\mathbf{R}) D_{r,r}^k(\mathbf{R}) \times \overline{D}_{l,m}^n(\mathbf{R}') \overline{D}_{r,r}^{k'}(\mathbf{R}') = \\ &= \frac{1}{8}e^{-a-c} \sum_{n,m,l,k,k',r,p,p'} (2n+1)(2k+1)(2k'+1)(2p+1)(2p'+1) i^{-r} e^{-im\omega} I_m\left(c\sqrt{1+\chi^2}\right) \times \\ &\quad \times I_r\left(\tau\left(1 - \frac{\lambda}{2\pi}\right)\right) \times \mathcal{L}_m^n(-a) \mathcal{L}_r^k(-bf) \mathcal{L}_r^{k'}(0) \times \begin{pmatrix} n & k & p \\ l & r & -l-r \end{pmatrix} \begin{pmatrix} n & k & p \\ m & r & -m-r \end{pmatrix} \times \\ &\quad \times \begin{pmatrix} n & k' & p' \\ l & r & -l-r \end{pmatrix} \begin{pmatrix} n & k' & p' \\ m & r & -m-r \end{pmatrix} \times D_{l+r,m+r}^p(\mathbf{R}) \overline{D}_{l+r,m+r}^{p'}(\mathbf{R}') \end{aligned} \quad (\text{B15})$$

Here  $E(\mathbf{R}, \mathbf{R}')$  is the local DNA energy term defined by eq. (6) in the main text, which is determined by the orientations of neighbour DNA segments described by rotation matrices  $\mathbf{R}$  and  $\mathbf{R}'$ , respectively.

By introducing new indexes  $s = l + r$  and  $t = m + r$  and switching the order of summation in eq. (B15), it is not very hard to obtain the following expansion formula:

$$e^{-E(\mathbf{R}, \mathbf{R}')} = \sum_{p,p',s,t} F_{pp'}^{st} D_{s,t}^p(\mathbf{R}) \overline{D}_{s,t}^{p'}(\mathbf{R}') \quad (\text{B16})$$

Where the expansion coefficients  $F_{pp'}^{st}$  are:

$$\begin{aligned} F_{pp'}^{st} &= \frac{1}{8} (2p+1)(2p'+1) e^{-a-c} \sum_{n,k,k',r} (2n+1)(2k+1)(2k'+1) i^{-r} e^{-i(t-r)\omega} I_{t-r}\left(c\sqrt{1+\chi^2}\right) I_r\left(\tau\left(1 - \frac{\lambda}{2\pi}\right)\right) \times \\ &\quad \times \mathcal{L}_{t-r}^n(-a) \mathcal{L}_r^k(-bf) \mathcal{L}_r^{k'}(0) \times \begin{pmatrix} n & k & p \\ s-r & r & -s \end{pmatrix} \begin{pmatrix} n & k & p \\ t-r & r & -t \end{pmatrix} \begin{pmatrix} n & k' & p' \\ s-r & r & -s \end{pmatrix} \begin{pmatrix} n & k' & p' \\ t-r & r & -t \end{pmatrix} \end{aligned} \quad (\text{B17})$$

To simplify the DNA partition function calculations, it will be more convenient to slightly re-organize eq. (B16) and (B17). First, it should be noted that from eq. (A9) it follows that functions  $D_{m,l}^n$  have the  $L_2$ -norm  $\|D_{m,l}^n\|_2 = \sqrt{\frac{8\pi^2}{2n+1}}$ :

$$\|D_{m,l}^n\|_2^2 = \int d\mathbf{R} \overline{D}_{m,l}^n(\mathbf{R}) D_{m,l}^n(\mathbf{R}) = \frac{8\pi^2}{2n+1} \quad (\text{B18})$$

Thus, while being orthogonal, the basis formed by  $D_{m,l}^n$  functions is not orthonormal. Using eq. (B18), we can easily normalize it by switching from  $D_{m,l}^n$  to  $\sqrt{\frac{2n+1}{8\pi^2}} D_{m,l}^n$  functions. By doing this, the earlier obtained expansion formula turns into:

$$e^{-E(\mathbf{R}, \mathbf{R}')} = \frac{1}{8\pi^2} \sum_{p,p',s,s',t,t'} \sqrt{(2p+1)(2p'+1)} \times T_{s,t,p}^{s',t',p'} D_{s,t}^p(\mathbf{R}) \overline{D}_{s',t'}^{p'}(\mathbf{R}') \quad (\text{B19})$$

Where the expansion coefficients  $T_{s,t,p}^{s',t',p'}$  are:

$$\begin{aligned} T_{s,t,p}^{s',t',p'} &= \delta_{ss'} \delta_{tt'} \times \pi^2 \sqrt{(2p+1)(2p'+1)} e^{-a-c} \sum_{n,k,k',r} (2n+1)(2k+1)(2k'+1) i^{-r} e^{-i(t-r)\omega} \times \\ &\times I_{t-r} \left( c\sqrt{1+\chi^2} \right) I_r \left( \tau \left( 1 - \frac{\lambda}{2\pi} \right) \right) \times \mathcal{L}_{t-r}^n(-a) \mathcal{L}_r^k(-bf) \mathcal{L}_r^{k'}(0) \times \\ &\times \begin{pmatrix} n & k & p \\ s-r & r & -s \end{pmatrix} \begin{pmatrix} n & k & p \\ t-r & r & -t \end{pmatrix} \begin{pmatrix} n & k' & p' \\ s-r & r & -s \end{pmatrix} \begin{pmatrix} n & k' & p' \\ t-r & r & -t \end{pmatrix} \end{aligned} \quad (\text{B20})$$

By putting  $s = s' = t = t' = 0$  in the above expression, we finally obtain eq. (10) from the main text describing the elements of the DNA transfer matrix in the case when the DNA boundary condition function has a symmetry about  $\mathbf{z}_0$ -axis of the lab coordinate system (see more details in Appendix C):

$$\begin{aligned} S_{pp'} &= T_{0,0,p}^{0,0,p'} = \pi^2 \sqrt{(2p+1)(2p'+1)} e^{-a-c} \sum_{n,k,k',r} (2n+1)(2k+1)(2k'+1) e^{ir(\omega - \frac{\pi}{2})} \times \\ &\times I_r \left( c\sqrt{1+\chi^2} \right) I_r \left( \tau \left( 1 - \frac{\lambda}{2\pi} \right) \right) \times \mathcal{L}_r^n(-a) \mathcal{L}_r^k(-bf) \mathcal{L}_r^{k'}(0) \times \begin{pmatrix} n & k & p \\ -r & r & 0 \end{pmatrix}^2 \begin{pmatrix} n & k' & p' \\ -r & r & 0 \end{pmatrix}^2 \end{aligned} \quad (\text{B21})$$

### Appendix C: DNA partition function

In this section, we will derive an analytic formula for the partition function of a mechanically stretched and twisted DNA using a semiflexible polymer model introduced in the main text. Representing DNA by a discretized polygonal chain consisting of short segments of equal length,  $b$ , and applying the Fuller's formula, the total conformational energy of the DNA,  $E_{\text{tot}}(\mathbf{R}_1, \dots, \mathbf{R}_N)$ , can be found as a sum of local energies  $E_j(\mathbf{R}_j, \mathbf{R}_{j+1})$  associated with the vertices connecting neighbouring DNA segments (eq. (5) in the main text):

$$E_{\text{tot}}(\mathbf{R}_1, \dots, \mathbf{R}_N) = \sum_{j=1}^{N-1} E_j(\mathbf{R}_j, \mathbf{R}_{j+1}) - bf(\mathbf{z}_0 \cdot \mathbf{R}_N \mathbf{z}_0) \quad (\text{C1})$$

As before,  $(\mathbf{R}_1, \dots, \mathbf{R}_j, \dots, \mathbf{R}_N)$  are Euler rotation matrices describing the orientations of individual DNA segments;  $N$  is the total number of DNA segments in the polygonal chain; and  $f = F/k_B T$ , where  $F$  is the stretching force applied to the DNA. Here and below all of the energies are in  $k_B T$  units. The rightmost term in eq. (C1) is the potential energy of the last DNA segment under load  $f$ , which is not accounted for by any of the local energies terms  $E_j(\mathbf{R}_j, \mathbf{R}_{j+1})$  defined as (eq. (6), main text):

$$\begin{aligned} E_j(\mathbf{R}_j, \mathbf{R}_{j+1}) &= \frac{a}{2} (\mathbf{R}_j \mathbf{z}_0 - \mathbf{R}_{j+1} \mathbf{z}_0)^2 + \frac{c}{2} (2\pi \Delta T w_j(\mathbf{R}_j, \mathbf{R}_{j+1}))^2 - bf(\mathbf{z}_0 \cdot \mathbf{R}_j \mathbf{z}_0) - \\ &\quad - \tau(2\pi - \lambda) \Delta L k_j^F(\mathbf{R}_j, \mathbf{R}_{j+1}) - \tau \lambda \Delta T w_j(\mathbf{R}_j, \mathbf{R}_{j+1}) \end{aligned} \quad (\text{C2})$$

Here  $a = A/b$  and  $c = C/b$  are the effective bending and twisting rigidities of the segments in the discretized model of DNA;  $\tau$  is the torque applied to DNA in  $k_B T$  units;  $\Delta L k_j^F(\mathbf{R}_j, \mathbf{R}_{j+1})$  and  $\Delta T w_j(\mathbf{R}_j, \mathbf{R}_{j+1})$  are the local contributions of the  $j^{\text{th}}$  and  $(j+1)^{\text{th}}$  DNA segments to the total change in the DNA linking and twist numbers, respectively;  $\lambda$  is the model parameter accounting for the repulsion energy associated with the DNA writhe formation [19].

Using eq. (C1)-(C2), the partition function of DNA under applied tension  $f$  and torque  $\tau$  can be found as:

$$Z_{f,\tau} = \int d\mathbf{R}_1 \dots d\mathbf{R}_N e^{-E_{\text{tot}}(\mathbf{R}_1, \dots, \mathbf{R}_N)} \xi(\mathbf{R}_N, \mathbf{R}_1) \quad (\text{C3})$$

Here  $\xi(\mathbf{R}_N, \mathbf{R}_1)$  is a function that imposes specific boundary conditions on the orientations of the DNA ends. To simplify eq. (C3), it is convenient to introduce the following notation:

$$T(\mathbf{R}, \mathbf{R}') = e^{-E(\mathbf{R}, \mathbf{R}')} \quad (\text{C4})$$

Where  $E(\mathbf{R}, \mathbf{R}')$  is the local energy contribution defined by eq. (C2) to the total energy of DNA by two neighbouring DNA segments, whose orientations in space are described by Euler matrices  $\mathbf{R}$  and  $\mathbf{R}'$  (i.e., in the case of  $\mathbf{R} = \mathbf{R}_j$  and  $\mathbf{R}' = \mathbf{R}_{j+1}$ :  $E(\mathbf{R}, \mathbf{R}') = E_j(\mathbf{R}_j, \mathbf{R}_{j+1})$ ). Substituting eq. (C4) and (C1) into eq. (C3), it is then easy to see that:

$$Z_{f,\tau} = \int d\mathbf{R}_1 \dots d\mathbf{R}_N \prod_{j=1}^{N-1} T(\mathbf{R}_j, \mathbf{R}_{j+1}) \times \sigma(\mathbf{R}_N, \mathbf{R}_1) \quad (\text{C5})$$

Where

$$\sigma(\mathbf{R}_N, \mathbf{R}_1) = e^{bf(\mathbf{z}_0 \cdot \mathbf{R}_N \mathbf{z}_0)} \xi(\mathbf{R}_N, \mathbf{R}_1) \quad (\text{C6})$$

Using eq. (B19)-(B20), it is possible to simplify eq. (C5), reducing all of the integrals to mere summations. Indeed, as was shown in Appendix B, each  $T(\mathbf{R}, \mathbf{R}')$  function in eq. (C5) can be expanded into the series of  $D_{m,l}^n$  functions as:

$$T(\mathbf{R}, \mathbf{R}') = \frac{1}{8\pi^2} \sum_{n,n',m,m',l,l'} \sqrt{(2n+1)(2n'+1)} \times T_{m,l,n}^{m',l',n'} D_{m,l}^n(\mathbf{R}) \bar{D}_{m',l'}^{n'}(\mathbf{R}') \quad (\text{C7})$$

Where the expansion coefficients  $T_{m,l,n}^{m',l',n'}$  are (eq. (B20)):

$$\begin{aligned} T_{m,l,n}^{m',l',n'} &= \delta_{mm'} \delta_{ll'} \times \pi^2 \sqrt{(2n+1)(2n'+1)} e^{-a-c} \sum_{p,k,k',r} (2p+1)(2k+1)(2k'+1) i^{-r} e^{-i(l-r)\omega} \times \\ &\times I_{l-r} \left( c \sqrt{1+\chi^2} \right) I_r \left( \tau \left( 1 - \frac{\lambda}{2\pi} \right) \right) \times \mathcal{L}_{l-r}^p(-a) \mathcal{L}_r^k(-bf) \mathcal{L}_r^{k'}(0) \times \\ &\times \begin{pmatrix} p & k & n \\ m-r & r & -m \end{pmatrix} \begin{pmatrix} p & k & n \\ l-r & r & -l \end{pmatrix} \begin{pmatrix} p & k' & n' \\ m-r & r & -m \end{pmatrix} \begin{pmatrix} p & k' & n' \\ l-r & r & -l \end{pmatrix} \end{aligned} \quad (\text{C8})$$

Here, as before,  $\chi = \frac{\tau\lambda}{2\pi c}$  and  $\omega = \tan^{-1}(\chi)$ .

Analogously, the boundary condition function  $\sigma(\mathbf{R}_N, \mathbf{R}_1)$  can be represented via  $D_{m,l}^n$ -series as:

$$\sigma(\mathbf{R}_N, \mathbf{R}_1) = \frac{1}{8\pi^2} \sum_{\substack{n_1, m_1, l_1 \\ n_N, m_N, l_N}} \sqrt{(2n_1+1)(2n_N+1)} \times \sigma_{m_N, l_N, n_N}^{m_1, l_1, n_1} D_{m_N, l_N}^{n_N}(\mathbf{R}_N) \bar{D}_{m_1, l_1}^{n_1}(\mathbf{R}_1) \quad (\text{C9})$$

Where the expansion coefficients  $\sigma_{m_N, l_N, n_N}^{m_1, l_1, n_1}$  can be found using the following expression:

$$\sigma_{m_N, l_N, n_N}^{m_1, l_1, n_1} = \frac{\sqrt{(2n_1+1)(2n_N+1)}}{8\pi^2} \int d\mathbf{R}_1 d\mathbf{R}_N \bar{D}_{m_N, l_N}^{n_N}(\mathbf{R}_N) \sigma(\mathbf{R}_N, \mathbf{R}_1) D_{m_1, l_1}^{n_1}(\mathbf{R}_1) \quad (\text{C10})$$

Substituting eq. (C7) and (C9) into eq. (C5) and using the orthogonality property of  $D_{m,l}^n$  functions (eq. (A9)), it is straightforward to show that:

$$Z_{f,\tau} = \sum_{\substack{m_1, \dots, m_N \\ l_1, \dots, l_N \\ n_1, \dots, n_N}} \left[ \prod_{j=1}^{N-1} T_{m_j, l_j, n_j}^{m_{j+1}, l_{j+1}, n_{j+1}} \times \sigma_{m_N, l_N, n_N}^{m_1, l_1, n_1} \right] \quad (\text{C11})$$

All that remains is to find the expansion coefficients  $\sigma_{m_N, l_N, n_N}^{m_1, l_1, n_1}$  in the above expression. While eq. (C9)-(C10) can be used to obtain the expansion series for any boundary condition function, here we will consider only two examples. First boundary condition function is obtained by assuming that the first and the last segments of DNA always stay collinear to the lab  $\mathbf{z}_0$ -axis during the DNA mechanical stretching and twisting. In this case, without loss of generality we can assume that the first DNA segment is fixed and has the same orientation as the lab coordinate system ( $\mathbf{R}_1 = \mathbf{I}$ ), while the last ( $N^{\text{th}}$ ) DNA segment can freely rotate about its  $\mathbf{z}_N$ -axis which always stays parallel to  $\mathbf{z}_0$ -axis of the fixed lab coordinate system ( $\mathbf{z}_N = \mathbf{R}_N \mathbf{z}_0 = \mathbf{z}_0$ ). Thus, we have:

$$\xi(\mathbf{R}_N, \mathbf{R}_1) = \delta(\mathbf{R}_1 - \mathbf{I}) \delta(\mathbf{R}_N \mathbf{z}_0 - \mathbf{z}_0) = \frac{\delta(\alpha_1) \delta(\beta_1) \delta(\gamma_1) \delta(\beta_N)}{2\pi \sin \beta_1 \sin \beta_N} \quad (\text{C12})$$

Here  $\delta(x)$  is the Dirac delta function.

From eq. (A1)-(A2), (A10), (C6), (C10) and (C12) it then follows that:

$$\begin{aligned} \sigma_{m_N, l_N, n_N}^{m_1, l_1, n_1} &= \frac{\sqrt{(2n_1+1)(2n_N+1)}}{8\pi^2} \int d\mathbf{R}_1 d\mathbf{R}_N \bar{D}_{m_N, l_N}^{n_N}(\mathbf{R}_N) e^{bf(\mathbf{z}_0 \cdot \mathbf{R}_N \mathbf{z}_0)} \delta(\mathbf{R}_1 - \mathbf{I}) \delta(\mathbf{R}_N \mathbf{z}_0 - \mathbf{z}_0) D_{m_1, l_1}^{n_1}(\mathbf{R}_1) = \\ &= \delta_{m_N 0} \delta_{l_N 0} \times \frac{\sqrt{(2n_1+1)(2n_N+1)}}{4\pi} e^{bf} \times P_{m_1, l_1}^{n_1}(1) \bar{P}_{m_N, l_N}^{n_N}(1) = \delta_{m_N 0} \delta_{l_N 0} \delta_{m_1 l_1} \times \frac{\sqrt{(2n_1+1)(2n_N+1)}}{4\pi} e^{bf} \end{aligned} \quad (\text{C13})$$



Another boundary condition function can be obtained by assuming that there is no constraint on the DNA ends' orientations (i.e., DNA ends can freely rotate). In this case:

$$\xi(\mathbf{R}_N, \mathbf{R}_1) = 1 \quad (\text{C14})$$

Then using eq. (A1), (A10), (B8), (C6) and (C10), we obtain the following expression for the expansion coefficients of the DNA ends free orientation boundary condition function:

$$\begin{aligned} \sigma_{m_N, l_N, n_N}^{m_1, l_1, n_1} &= \frac{\sqrt{(2n_1+1)(2n_N+1)}}{8\pi^2} \int d\mathbf{R}_1 d\mathbf{R}_N \bar{D}_{m_N, l_N}^{n_N}(\mathbf{R}_N) e^{bf(\mathbf{z}_0 \cdot \mathbf{R}_N \mathbf{z}_0)} D_{m_1, l_1}^{n_1}(\mathbf{R}_1) = \\ &= \delta_{m_1 0} \delta_{l_1 0} \delta_{m_N 0} \delta_{l_N 0} \times 2\pi^2 \sqrt{(2n_1+1)(2n_N+1)} \times \mathcal{L}_0^{n_1}(0) \mathcal{L}_0^{n_N}(-bf) \end{aligned} \quad (\text{C15})$$

From eq. (C13) and (C15) it can be seen that, despite slightly different forms, the expansion coefficients of the two boundary condition functions considered here have the same term:  $\delta_{m_N 0} \delta_{l_N 0}$ . Together with  $\delta_{m m'} \delta_{l l'}$  term in eq. (C8) it leads to a peculiar domino-like effect resulting in nullification of all of the tesseral harmonics in the final expression for the DNA partition function (nullification of all of the terms with non-zero indexes  $m$  and  $l$ ). I.e., by substituting eq. (C13) or (C15) into eq. (C11), we have:

$$Z_{f, \tau} = \sum_{n_1, \dots, n_N} \left[ \prod_{j=1}^{N-1} T_{0,0,n_j}^{0,0,n_{j+1}} \times \sigma_{0,0,n_N}^{0,0,n_1} \right] \quad (\text{C16})$$

In fact, it can be proved that eq. (C16) holds in a much more general case when the boundary condition function,  $\xi(\mathbf{R}_N, \mathbf{R}_1)$ , possesses a symmetry about  $\mathbf{z}_0$ -axis of the lab coordinate system with respect to either the first or the last DNA segments' orientations, i.e., when rotation of one of these DNA segments about  $\mathbf{z}_0$ -axis does not change the value of the boundary condition function. It is clear that the above two examples of the DNA boundary condition function, namely, eq. (C12) and (C14), fall into this category.

To streamline eq. (C16), it is convenient to introduce two square matrices,  $\mathbf{S}$  and  $\mathbf{V}$ , whose elements are defined as:

$$S_{nn'} = T_{0,0,n}^{0,0,n'} \quad \text{and} \quad V_{nn'} = \sigma_{0,0,n}^{0,0,n'} \quad (\text{C17})$$

Then eq. (C16) takes a very simple and compact form:

$$Z_{f, \tau} = \text{Tr}(\mathbf{S}^{N-1} \mathbf{V}) \quad (\text{C18})$$

It should be noted that while in the above expression square matrices  $\mathbf{S}$  and  $\mathbf{V}$  in the general case have infinite sizes, calculations show that the value of the DNA partition function is typically determined by first  $M \approx 10$ -15 harmonics [29, 40]. Thus, in real computations it makes sense to use finite  $M \times M$  square matrices  $\mathbf{S}$  and  $\mathbf{V}$  in eq. (C18), which include only first  $M$  rows and columns.

In conclusion, to obtain the partition function of a mechanically stretched and twisted DNA, one has to:

- 1) calculate the elements  $S_{nn'}$  and  $V_{nn'}$  of matrices  $\mathbf{S}$  and  $\mathbf{V}$  for  $0 \leq n, n' \leq M-1$ :

$$\begin{aligned} S_{nn'} &= \pi^2 \sqrt{(2n+1)(2n'+1)} e^{-a-c} \sum_{p,k,k',r} (2p+1)(2k+1)(2k'+1) e^{ir(\omega - \frac{\pi}{2})} \times \\ &\quad \times I_r\left(c\sqrt{1+\chi^2}\right) I_r\left(\tau\left(1 - \frac{\lambda}{2\pi}\right)\right) \times \mathcal{L}_r^p(-a) \mathcal{L}_r^k(-bf) \mathcal{L}_r^{k'}(0) \times \begin{pmatrix} p & k & n \\ -r & r & 0 \end{pmatrix}^2 \begin{pmatrix} p & k' & n' \\ -r & r & 0 \end{pmatrix}^2 \end{aligned} \quad (\text{C19})$$

and

$$\begin{cases} V_{nn'} = \frac{1}{4\pi} e^{bf} \sqrt{(2n+1)(2n'+1)} & \text{- the DNA ends } \mathbf{z}_0\text{-axis collinear boundary condition} \\ V_{nn'} = 2\pi^2 \mathcal{L}_0^{n'}(0) \mathcal{L}_0^n(-bf) \sqrt{(2n+1)(2n'+1)} = & \text{- the DNA ends free orientation boundary condition,} \\ = 8\pi^2 \delta_{n'0} i_n(bf) \sqrt{2n+1} & \text{see formula (D3) for } \mathcal{L}_0^n \text{ functions in Appendix D} \end{cases}$$

- 2) compute the trace of  $\mathbf{S}$  and  $\mathbf{V}$  matrices product according to eq. (C18) to get the value of the DNA partition function.

### Appendix D: Consistency of the semiflexible polymer model with the worm-like chain model of DNA

In order to show that the semiflexible polymer model described in Appendix C is consistent with the classical worm-like chain model of a torsionally relaxed DNA (when  $\tau = 0$ ) [41] we will require several simple relations between the special functions introduced in Appendix A.

First, from the definition of polynomials  $P_{m,l}^n$  it can be shown that in the special case of  $m = l = 0$  they converge to very well-known Legendre polynomials,  $P_n$ :

$$P_{0,0}^n(x) = P_n(x) \quad (\text{D1})$$

Furthermore, by applying inverse Fourier transform to Legendre polynomials, we get (p. 770, ref. [39]):

$$j_n(x) = \frac{1}{2i^n} \int_{-1}^1 P_n(\omega) e^{i\omega x} d\omega \quad (\text{D2})$$

Where  $j_n(x)$  are spherical Bessel functions of the first kind and  $i$  is imaginary unit.

Substituting eq. (D1)-(D2) into eq. (B8), it is then easy to find that:

$$\mathcal{L}_0^n(s) = \int_{-1}^1 P_{0,0}^n(x) e^{-sx} dx = \int_{-1}^1 P_n(x) e^{-sx} dx = 2i^n j_n(is) = 2i_n(-s) \quad (\text{D3})$$

Where  $i_n(x) = i^{-n} j_n(ix)$  is the modified spherical Bessel function of the first kind, which has the parity of  $(-1)^n$  (i.e.,  $i_n(-x) = (-1)^n i_n(x)$ , see p. 735-736 in ref. [39]).

The values of modified Bessel functions of the first kind,  $I_n(x)$ , as well as spherical modified Bessel functions of the first kind,  $i_n(x)$ , at point  $x = 0$  are (p. 714 and p. 735, ref. [39]):

$$I_n(0) = i_n(0) = \delta_{n0} \quad (\text{D4})$$

Finally, it can be shown that in the case of  $n_2 = m_1 = m_2 = m_3 = 0$  Wigner 3-j symbols equal to (p. 1059, ref. [42]):

$$\begin{pmatrix} n_1 & 0 & n_3 \\ 0 & 0 & 0 \end{pmatrix} = \frac{(-1)^{n_1}}{\sqrt{2n_1+1}} \delta_{n_1 n_3} \quad (\text{D5})$$

Let's now use the above formulas to find out the values of the transfer matrix elements,  $S_{nn'}$ , for a torsionally relaxed DNA ( $\tau = 0$ ). From eq. (C19), we have:

$$\begin{aligned} S_{nn'}|_{\tau=0} &= \left[ \tau = 0 \Rightarrow \chi = \frac{\tau\lambda}{2\pi c} = 0, \omega = \tan^{-1}(\chi) = 0 \right] = \\ &= \pi^2 \sqrt{(2n+1)(2n'+1)} e^{-a-c} \sum_{p,k,k',r} (2p+1)(2k+1)(2k'+1) i^{-r} I_r(c) I_r(0) \mathcal{L}_r^p(-a) \mathcal{L}_r^k(-bf) \mathcal{L}_r^{k'}(0) \times \\ &\quad \times \begin{pmatrix} p & k & n \\ -r & r & 0 \end{pmatrix}^2 \begin{pmatrix} p & k' & n' \\ -r & r & 0 \end{pmatrix}^2 = [\text{eq. (D4)} \Rightarrow r = 0] = \\ &= \pi^2 \sqrt{(2n+1)(2n'+1)} e^{-a-c} I_0(c) \sum_{p,k,k'} (2p+1)(2k+1)(2k'+1) \mathcal{L}_0^p(-a) \mathcal{L}_0^k(-bf) \mathcal{L}_0^{k'}(0) \times \\ &\quad \times \begin{pmatrix} p & k & n \\ 0 & 0 & 0 \end{pmatrix}^2 \begin{pmatrix} p & k' & n' \\ 0 & 0 & 0 \end{pmatrix}^2 = [\text{eq. (D3) and (D4)} \Rightarrow k' = 0] = \\ &= 2\pi^2 \sqrt{(2n+1)(2n'+1)} e^{-a-c} I_0(c) \sum_{p,k} (2p+1)(2k+1) \mathcal{L}_0^p(-a) \mathcal{L}_0^k(-bf) \begin{pmatrix} p & k & n \\ 0 & 0 & 0 \end{pmatrix}^2 \begin{pmatrix} p & 0 & n' \\ 0 & 0 & 0 \end{pmatrix}^2 = \\ &= [\text{eq. (D5)} \Rightarrow p = n'] = 2\pi^2 \sqrt{(2n+1)(2n'+1)} e^{-a-c} I_0(c) \mathcal{L}_0^{n'}(-a) \sum_k (2k+1) \mathcal{L}_0^k(-bf) \begin{pmatrix} n' & k & n \\ 0 & 0 & 0 \end{pmatrix}^2 = \\ &= [\text{eq. (D3)}] = 2\pi e^{-c} I_0(c) \times 4\pi \sqrt{(2n+1)(2n'+1)} e^{-a} i_{n'}(a) \sum_k (2k+1) i_k(bf) \begin{pmatrix} n' & k & n \\ 0 & 0 & 0 \end{pmatrix}^2 \quad (\text{D6}) \end{aligned}$$

Comparing the final result of eq. (D6) to the elements of the DNA transfer matrix in the WLC model (eq. 7 from ref. [30] where  $m = m' = 0$ ), we see that the only difference between the two models is the coefficient  $2\pi e^{-c} I_0(c)$ . This term, however, does not have any influence on the equilibrium DNA behaviour since it simply results in the DNA free energy offset by a fixed constant (namely, by  $(N-1) \times \ln[2\pi e^{-c} I_0(c)]$ ). Thus, it can be concluded that the semiflexible polymer model of DNA indeed converges to the WLC model in the case of a torsionally relaxed DNA.

### Appendix E: DNA structural transitions

Experimental studies show that depending on the value of the applied force and torque DNA can experience transitions between several alternative structural states named B-DNA, L-DNA, P-DNA, etc., which have very different geometric and elastic characteristics. During recent years most of these DNA states were studied and described in great details in a number of single-molecule experiments. However, despite all of the gained knowledge, it is still an open question how the physical properties of different DNA structures and their sensitivity to the nucleotide sequence determine the probability to find various DNA segments in one or the other structural state, and how this probability is modulated by the mechanical constraints applied to DNA as well as other environmental factors, such as solution ionic strength or presence of DNA-binding proteins.

To address these questions, in this section we are going to extend the semiflexible polymer model and transfer-matrix technique developed in the preceding Appendices by taking into account that DNA segments can transit between several alternative structural states (B-DNA, L-DNA, P-DNA). For this purpose, we will first consider a simpler abstract case when each DNA segment can switch between only two alternative states and then generalize the obtained formulas for an arbitrary number of possible DNA segments' structures.

Let's consider a DNA polymer whose segments can transit between two alternative structures, which we enumerate by 0 and 1. In this case, the global configuration of the DNA is completely determined by the two sets: 1) collection of Euler rotation matrices,  $(\mathbf{R}_1, \dots, \mathbf{R}_N)$ , describing the orientation of each DNA segment; and 2) set  $(k_1, \dots, k_N)$  indicating the structural state of each of the DNA segments (for any segment  $j : k_j = 0$  or 1). The total energy of the DNA,  $E_{\text{tot}}$ , then can be found as:

$$E_{\text{tot}}(k_1, \dots, k_N, \mathbf{R}_1, \dots, \mathbf{R}_N) = \sum_{j=1}^{N-1} E_j^{(k_j)}(\mathbf{R}_j, \mathbf{R}_{j+1}) + q \left( \mu - 2\pi\tau \Delta l k_0^{(1)} \right) \sum_{j=1}^N \delta_{k_{j1}} - b_{k_N} f(\mathbf{z}_0 \cdot \mathbf{R}_N \mathbf{z}_0) \quad (\text{E1})$$

Where  $b_{k_j}$  is the length of the  $j^{\text{th}}$  DNA segment in the respective state (i.e.,  $b_{k_j} = b_0$  or  $b_{k_j} = b_1$ );  $q$  is the number of base-pairs in each of the DNA segments (assumed to be the same for both states 0 and 1);  $\mu$  is the free energy difference between the two DNA states calculated per single base-pair (i.e., energy price per single base-pair required for the DNA transition from state 0 to state 1), and  $\Delta l k_0^{(1)} = l k_{0,1} - l k_{0,0}$  is the difference between the relaxed linking numbers of DNA in state 1 ( $l k_{0,1}$ ) and state 0 ( $l k_{0,0}$ ) calculated per single DNA base-pair. As before,  $E_j^{(k_j)}(\mathbf{R}_j, \mathbf{R}_{j+1})$  is the local contribution of the vertex connecting neighbour segments to the total DNA energy in the case when the  $j^{\text{th}}$  DNA segment is in state  $k_j$ :

$$E_j^{(k_j)}(\mathbf{R}_j, \mathbf{R}_{j+1}) = \frac{a_{k_j}}{2} (\mathbf{R}_j \mathbf{z}_0 - \mathbf{R}_{j+1} \mathbf{z}_0)^2 + \frac{c_{k_j}}{2} (2\pi \Delta T w_j(\mathbf{R}_j, \mathbf{R}_{j+1}))^2 - b_{k_j} f(\mathbf{z}_0 \cdot \mathbf{R}_j \mathbf{z}_0) - \tau (2\pi - \lambda_{k_j}) \Delta L k_j^F(\mathbf{R}_j, \mathbf{R}_{j+1}) - \tau \lambda_{k_j} \Delta T w_j(\mathbf{R}_j, \mathbf{R}_{j+1}) \quad (\text{E2})$$

Here  $a_{k_j}$ ,  $c_{k_j}$  and  $\lambda_{k_j}$  are the model parameters describing the bending and twisting rigidities of the  $j^{\text{th}}$  DNA segment as well as the DNA writhe repulsion energy ( $\lambda$  parameter) in the respective states (i.e.,  $a_{k_j} = a_0$  or  $a_{k_j} = a_1$ ,  $c_{k_j} = c_0$  or  $c_{k_j} = c_1$ , and  $\lambda_{k_j} = \lambda_0$  or  $\lambda_{k_j} = \lambda_1$ ).

Using eq. (E1), the partition function of DNA can be calculated as:

$$Z_{f,\tau} = \sum_{k_1, \dots, k_N=0}^1 \int d\mathbf{R}_1 \dots d\mathbf{R}_N e^{-E_{\text{tot}}(k_1, \dots, k_N, \mathbf{R}_1, \dots, \mathbf{R}_N)} \xi(\mathbf{R}_N, \mathbf{R}_1) \quad (\text{E3})$$

Where, as before,  $\xi(\mathbf{R}_N, \mathbf{R}_1)$  is a function that imposes specific boundary conditions on the orientations of the DNA first and last segments. Similarly to Appendix C, eq. (E3) can be simplified by introducing the following notation:

$$T_k(\mathbf{R}, \mathbf{R}') = e^{-E^{(k)}(\mathbf{R}, \mathbf{R}') - q(\mu - 2\pi\tau \Delta l k_0^{(1)}) \delta_{k1}} \quad (\text{E4})$$

Substituting eq. (E4) and (E1) into eq. (E3) we get:

$$Z_{f,\tau} = \sum_{k_1, \dots, k_N=0}^1 \int d\mathbf{R}_1 \dots d\mathbf{R}_N \prod_{j=1}^{N-1} T_{k_j}(\mathbf{R}_j, \mathbf{R}_{j+1}) \times \sigma_{k_N}(\mathbf{R}_N, \mathbf{R}_1) \quad (\text{E5})$$

Where

$$\sigma_{k_N}(\mathbf{R}_N, \mathbf{R}_1) = e^{-q(\mu - 2\pi\tau\Delta lk_0^{(1)})\delta_{k_N 1} + b_{k_N} f(\mathbf{z}_0 \cdot \mathbf{R}_N \mathbf{z}_0)} \xi(\mathbf{R}_N, \mathbf{R}_1) \quad (\text{E6})$$

It should be noted that the DNA partition function defined by eq. (E5) obeys a set of important recurrence relations, which can be used to further simplify it. To derive them, let's first define intermediary partition functions as:

$$Z_s(k_s, \mathbf{R}_s, \mathbf{R}_1) = \sum_{k_{s+1}, \dots, k_N=0}^1 \int d\mathbf{R}_{s+1} \dots d\mathbf{R}_N \prod_{j=s}^{N-1} T_{k_j}(\mathbf{R}_j, \mathbf{R}_{j+1}) \times \sigma_{k_N}(\mathbf{R}_N, \mathbf{R}_1) \quad (\text{E7})$$

Here  $1 \leq s \leq N-1$ . Then from eq. (E5) and (E7) it is not hard to see that:

$$Z_{f,\tau} = \int d\mathbf{R}_1 [Z_1(0, \mathbf{R}_1, \mathbf{R}_1) + Z_1(1, \mathbf{R}_1, \mathbf{R}_1)] = \int d\mathbf{R}_1 \begin{pmatrix} 1 & 1 \end{pmatrix} \times \begin{pmatrix} Z_1(0, \mathbf{R}_1, \mathbf{R}_1) \\ Z_1(1, \mathbf{R}_1, \mathbf{R}_1) \end{pmatrix} \quad (\text{E8})$$

Furthermore, from the definition of the intermediary partition functions it follows that they obey the following recurrence relation:

$$Z_{s-1}(k_{s-1}, \mathbf{R}_{s-1}, \mathbf{R}_1) = \int d\mathbf{R}_s (T_{k_{s-1}}(\mathbf{R}_{s-1}, \mathbf{R}_s) \begin{pmatrix} Z_s(0, \mathbf{R}_s, \mathbf{R}_1) \\ Z_s(1, \mathbf{R}_s, \mathbf{R}_1) \end{pmatrix}) \quad (\text{E9})$$

By using the vector-valued integration technique (see Appendix F for more details), eq. (E9) can be re-written as:

$$\begin{pmatrix} Z_{s-1}(0, \mathbf{R}_{s-1}, \mathbf{R}_1) \\ Z_{s-1}(1, \mathbf{R}_{s-1}, \mathbf{R}_1) \end{pmatrix} = \int d\mathbf{R}_s \mathbf{T}(\mathbf{R}_{s-1}, \mathbf{R}_s) \times \begin{pmatrix} Z_s(0, \mathbf{R}_s, \mathbf{R}_1) \\ Z_s(1, \mathbf{R}_s, \mathbf{R}_1) \end{pmatrix} \quad (\text{E10})$$

Where the transfer matrix  $\mathbf{T}(\mathbf{R}_{s-1}, \mathbf{R}_s)$  is:

$$\mathbf{T}(\mathbf{R}_{s-1}, \mathbf{R}_s) = \begin{pmatrix} T_0(\mathbf{R}_{s-1}, \mathbf{R}_s) & T_0(\mathbf{R}_{s-1}, \mathbf{R}_s) \\ T_1(\mathbf{R}_{s-1}, \mathbf{R}_s) & T_1(\mathbf{R}_{s-1}, \mathbf{R}_s) \end{pmatrix} \quad (\text{E11})$$

Combining eq. (E7)-(E8) and (E10), we obtain the following expression for the DNA partition function:

$$Z_{f,\tau} = \begin{pmatrix} 1 & 1 \end{pmatrix} \int d\mathbf{R}_1 \dots d\mathbf{R}_N \prod_{j=1}^{N-1} \mathbf{T}(\mathbf{R}_j, \mathbf{R}_{j+1}) \times \boldsymbol{\sigma}(\mathbf{R}_N, \mathbf{R}_1) \quad (\text{E12})$$

Where  $\boldsymbol{\sigma}(\mathbf{R}_N, \mathbf{R}_1)$  matrix is defined as:

$$\boldsymbol{\sigma}(\mathbf{R}_N, \mathbf{R}_1) = \begin{pmatrix} \sigma_0(\mathbf{R}_N, \mathbf{R}_1) \\ \sigma_1(\mathbf{R}_N, \mathbf{R}_1) \end{pmatrix} = \xi(\mathbf{R}_N, \mathbf{R}_1) \begin{pmatrix} e^{b_0 f(\mathbf{z}_0 \cdot \mathbf{R}_N \mathbf{z}_0)} \\ e^{-q(\mu - 2\pi\tau\Delta lk_0^{(1)})\delta_{k_1} + b_1 f(\mathbf{z}_0 \cdot \mathbf{R}_N \mathbf{z}_0)} \end{pmatrix} \quad (\text{E13})$$

It is interesting to note that eq. (E12) determining the partition function of DNA has the same functional form as eq. (C5) obtained for a simpler case when all of the DNA segments have a fixed structural state. The rest of the equations in this section practically follow the steps from Appendix C. Indeed, from eq. (E2) and (E4) it is not hard to see that functions  $T_k(\mathbf{R}, \mathbf{R}')$  ( $k = 0, 1$ ) have the same form as in eq. (C4) (up to a constant multiplier  $e^{-q(\mu - 2\pi\tau\Delta lk_0^{(1)})\delta_{k_1}}$ ). Therefore, they can be expanded into the  $D_{m,l}^n$ -series using the formulas, which are similar to eq. (C7)-(C8):

$$T_k(\mathbf{R}, \mathbf{R}') = \frac{1}{8\pi^2} \sum_{n, n', m, m', l, l'} \sqrt{(2n+1)(2n'+1)} \times (T_k)_{m, l, n}^{m', l', n'} D_{m, l}^n(\mathbf{R}) \overline{D}_{m', l'}^{n'}(\mathbf{R}') \quad (\text{E14})$$

Where the expansion coefficients  $(T_k)_{m,l,n}^{m',l',n'}$  are:

$$\begin{aligned} (T_k)_{m,l,n}^{m',l',n'} &= \delta_{mm'} \delta_{ll'} \times e^{-q(\mu-2\pi\tau\Delta lk_0^{(1)})\delta_{k1}} \times \pi^2 \sqrt{(2n+1)(2n'+1)} e^{-a_k - c_k} \sum_{p,s,s',r} (2p+1)(2s+1)(2s'+1) \times \\ &\times i^{-r} e^{-i(l-r)\omega_k} \times I_{l-r} \left( c_k \sqrt{1+\chi_k^2} \right) I_r \left( \tau \left( 1 - \frac{\lambda_k}{2\pi} \right) \right) \times \mathcal{L}_{l-r}^p(-a_k) \mathcal{L}_r^s(-b_k f) \mathcal{L}_r^{s'}(0) \times \\ &\times \begin{pmatrix} p & s & n \\ m-r & r & -m \end{pmatrix} \begin{pmatrix} p & s & n \\ l-r & r & -l \end{pmatrix} \begin{pmatrix} p & s' & n' \\ m-r & r & -m \end{pmatrix} \begin{pmatrix} p & s' & n' \\ l-r & r & -l \end{pmatrix} \end{aligned} \quad (\text{E15})$$

Here  $\chi_k = \frac{\tau\lambda_k}{2\pi c_k}$  and  $\omega_k = \tan^{-1}(\chi_k)$ .

Taking into account the linear property of matrices, it is then clear that the transfer matrix  $\mathbf{T}(\mathbf{R}, \mathbf{R}')$  defined by eq. (E11) can be expanded into the  $D_{m,l}^n$ -series as well:

$$\mathbf{T}(\mathbf{R}, \mathbf{R}') = \frac{1}{8\pi^2} \sum_{n,n',m,m',l,l'} \sqrt{(2n+1)(2n'+1)} \times \mathbf{T}_{m,l,n}^{m',l',n'} D_{m,l}^n(\mathbf{R}) \overline{D}_{m',l'}^{n'}(\mathbf{R}') \quad (\text{E16})$$

Where

$$\mathbf{T}_{m,l,n}^{m',l',n'} = \begin{pmatrix} (T_0)_{m,l,n}^{m',l',n'} & (T_0)_{m,l,n}^{m',l',n'} \\ (T_1)_{m,l,n}^{m',l',n'} & (T_1)_{m,l,n}^{m',l',n'} \end{pmatrix} \quad (\text{E17})$$

Analogously, for the boundary condition vector we have:

$$\boldsymbol{\sigma}(\mathbf{R}_N, \mathbf{R}_1) = \frac{1}{8\pi^2} \sum_{\substack{n_1, m_1, l_1 \\ n_N, m_N, l_N}} \sqrt{(2n_1+1)(2n_N+1)} \times \boldsymbol{\sigma}_{m_N, l_N, n_N}^{m_1, l_1, n_1} D_{m_N, l_N}^{n_N}(\mathbf{R}_N) \overline{D}_{m_1, l_1}^{n_1}(\mathbf{R}_1) \quad (\text{E18})$$

Where the vector  $\boldsymbol{\sigma}_{m_N, l_N, n_N}^{m_1, l_1, n_1}$  of the expansion coefficients is defined as:

$$\boldsymbol{\sigma}_{m_N, l_N, n_N}^{m_1, l_1, n_1} = \begin{pmatrix} (\sigma_0)_{m_N, l_N, n_N}^{m_1, l_1, n_1} \\ (\sigma_1)_{m_N, l_N, n_N}^{m_1, l_1, n_1} \end{pmatrix} \quad (\text{E19})$$

Comparing eq. (E13) and (E18)-(E19) to eq. (C9)-(C10), (C12)-(C15), and taking into account eq. (D3)-(D4), it is easy to see that in the case of the DNA ends  $\mathbf{z}_0$ -axis collinear boundary condition:

$$\frac{(\sigma_0)_{m_N, l_N, n_N}^{m_1, l_1, n_1}}{e^{b_0 f}} = \frac{(\sigma_1)_{m_N, l_N, n_N}^{m_1, l_1, n_1}}{e^{-q(\mu-2\pi\tau\Delta lk_0^{(1)})+b_1 f}} = \delta_{m_N 0} \delta_{l_N 0} \delta_{m_1 l_1} \times \frac{\sqrt{(2n_1+1)(2n_N+1)}}{4\pi} \quad (\text{E20})$$

While for the DNA ends free orientation boundary condition we have:

$$\frac{(\sigma_0)_{m_N, l_N, n_N}^{m_1, l_1, n_1}}{i_{n_N}(b_0 f)} = \frac{(\sigma_1)_{m_N, l_N, n_N}^{m_1, l_1, n_1}}{e^{-q(\mu-2\pi\tau\Delta lk_0^{(1)})} i_{n_N}(b_1 f)} = \delta_{n_1 0} \delta_{m_1 0} \delta_{l_1 0} \delta_{m_N 0} \delta_{l_N 0} \times 8\pi^2 \sqrt{2n_N+1} \quad (\text{E21})$$

Substituting eq. (E16) and (E18) into eq. (E12), and applying the orthogonality property of  $D_{m,l}^n$  functions (eq. (A9)), we can reduce all of the integrals in eq. (E12) to mere summations over the indexes of the expansion coefficients' matrices:

$$Z_{f,\tau} = (1 \ 1) \sum_{\substack{m_1, \dots, m_N \\ l_1, \dots, l_N \\ n_1, \dots, n_N}} \left[ \prod_{j=1}^{N-1} \mathbf{T}_{m_j, l_j, n_j}^{m_{j+1}, l_{j+1}, n_{j+1}} \times \boldsymbol{\sigma}_{m_N, l_N, n_N}^{m_1, l_1, n_1} \right] \quad (\text{E22})$$

Similarly to Appendix C, the presence of  $\delta$ -terms in eq. (E15) and (E20)-(E21) leads to complete disappearance of all tesseral harmonics in the expression for the DNA partition function:

$$Z_{f,\tau} = (1 \ 1) \sum_{n_1, \dots, n_N} \left[ \prod_{j=1}^{N-1} \mathbf{T}_{0,0,n_j}^{0,0,n_{j+1}} \times \boldsymbol{\sigma}_{0,0,n_N}^{0,0,n_1} \right] \quad (\text{E23})$$

To further simplify eq. (E23), we note that while the above matrices  $\mathbf{T}_{0,0,n_j}^{0,0,n_{j+1}}$  and  $\sigma_{0,0,n_N}^{0,0,n_1}$  have an infinite size, calculations show that the value of the DNA partition function is mainly determined only by first  $M \approx 10-15$  harmonics [29, 40]. Thus, to compute the DNA partition function, it makes sense to introduce finite square matrices  $\mathbf{S}_0, \mathbf{S}_1, \mathbf{V}_0$  and  $\mathbf{V}_1$  of  $M \times M$  size comprising only the first  $M$  rows/columns of the expansion coefficients defined by eq. (E14)-(E21), i.e.:

$$(S_0)_{nn'} = (T_0)_{0,0,n}^{0,0,n'}, \quad (S_1)_{nn'} = (T_1)_{0,0,n}^{0,0,n'}, \quad (V_0)_{nn'} = (\sigma_0)_{0,0,n}^{0,0,n'}, \quad \text{and} \quad (V_1)_{nn'} = (\sigma_1)_{0,0,n}^{0,0,n'} \quad (\text{E24})$$

Substituting eq. (E17), (E19) and (E24) into eq. (E23), it is not hard to see that all of the sums over indexes  $n_2, \dots, n_{N-1}$  have the following form:

$$\sum_t \begin{pmatrix} (S_0)_{st} & (S_0)_{st} \\ (S_1)_{st} & (S_1)_{st} \end{pmatrix} \begin{pmatrix} (S_0)_{tp} & (S_0)_{tp} \\ (S_1)_{tp} & (S_1)_{tp} \end{pmatrix} = \begin{pmatrix} (\mathbf{S}_0^2 + \mathbf{S}_0\mathbf{S}_1)_{sp} & (\mathbf{S}_0^2 + \mathbf{S}_0\mathbf{S}_1)_{sp} \\ (\mathbf{S}_1^2 + \mathbf{S}_1\mathbf{S}_0)_{sp} & (\mathbf{S}_1^2 + \mathbf{S}_1\mathbf{S}_0)_{sp} \end{pmatrix} \quad (\text{E25})$$

For the sake of the formulas simplicity, let's introduce a new notation:

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \Big|_{st} = \begin{pmatrix} A_{st} & B_{st} \\ C_{st} & D_{st} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix} \Big|_{st} = \begin{pmatrix} A_{st} \\ B_{st} \end{pmatrix} \quad (\text{E26})$$

Where  $\mathbf{A}, \mathbf{B}, \mathbf{C}$  and  $\mathbf{D}$  are arbitrary square matrices of the same size. Then eq. (E25) can be re-written as:

$$\sum_t \begin{pmatrix} (S_0)_{st} & (S_0)_{st} \\ (S_1)_{st} & (S_1)_{st} \end{pmatrix} \begin{pmatrix} (S_0)_{tp} & (S_0)_{tp} \\ (S_1)_{tp} & (S_1)_{tp} \end{pmatrix} = \sum_t \left[ \begin{pmatrix} \mathbf{S}_0 & \mathbf{S}_0 \\ \mathbf{S}_1 & \mathbf{S}_1 \end{pmatrix} \Big|_{st} \times \begin{pmatrix} \mathbf{S}_0 & \mathbf{S}_0 \\ \mathbf{S}_1 & \mathbf{S}_1 \end{pmatrix} \Big|_{tp} \right] = \begin{pmatrix} \mathbf{S}_0 & \mathbf{S}_0 \\ \mathbf{S}_1 & \mathbf{S}_1 \end{pmatrix}^2 \Big|_{sp} \quad (\text{E27})$$

I.e., eq. (E27) in some sense is a generalization of the matrices multiplication formula. Applying it to eq. (E23)  $(N-2)$  times, it can be shown that the expression for the DNA partition function reduces to:

$$Z_{f,\tau} = (1 \ 1) \sum_{n_1, n_N} \left[ \hat{\mathbf{S}}^{N-1} \Big|_{n_1 n_N} \times \hat{\mathbf{V}} \Big|_{n_N n_1} \right] \quad (\text{E28})$$

Where block-matrices  $\hat{\mathbf{S}}$  and  $\hat{\mathbf{V}}$  are:

$$\hat{\mathbf{S}} = \begin{pmatrix} \mathbf{S}_0 & \mathbf{S}_0 \\ \mathbf{S}_1 & \mathbf{S}_1 \end{pmatrix} \quad \text{and} \quad \hat{\mathbf{V}} = \begin{pmatrix} \mathbf{V}_0 \\ \mathbf{V}_1 \end{pmatrix} \quad (\text{E29})$$

Finally, after a few simple algebraic re-arrangements, eq. (E28) can be presented in the following form:

$$Z_{f,\tau} = \text{Tr}(\mathbf{U} \hat{\mathbf{S}}^{N-1} \hat{\mathbf{V}}) \quad (\text{E30})$$

Where block-matrix  $\mathbf{U} = (\mathbf{I} \ \mathbf{I})$ , and  $\mathbf{I}$  is the square  $M \times M$  identity matrix ( $I_{ml} = \delta_{ml}$ ).

While eq. (E24)-(E30) were obtained under the assumption of the finiteness of matrices  $\mathbf{S}_0, \mathbf{S}_1, \mathbf{V}_0$  and  $\mathbf{V}_1$ , for experienced readers it will not take too much efforts to show validity of these formulas in the general case of infinite matrices albeit with slightly different definitions of matrices  $\hat{\mathbf{S}}$  and  $\hat{\mathbf{V}}$  (for example, they can be represented as alternating sequences of  $\mathbf{S}_0, \mathbf{S}_1$  and  $\mathbf{V}_0, \mathbf{V}_1$  matrices' rows and columns).

With the help of the methodology described above it is then not very hard to prove very similar results for the general case when the DNA segments can switch between an arbitrary number,  $K$ , of structural states. It can be shown that in this case the DNA partition function still can be calculated via eq. (E30). However, this time the transfer block-matrix,  $\hat{\mathbf{S}}$ , will have the following form:

$$\hat{\mathbf{S}} = \begin{pmatrix} \mathbf{S}_0 & \mathbf{S}_0 & \cdots & \mathbf{S}_0 \\ \mathbf{S}_1 & \mathbf{S}_1 & \cdots & \mathbf{S}_1 \\ \vdots & \vdots & & \vdots \\ \mathbf{S}_{K-1} & \mathbf{S}_{K-1} & \cdots & \mathbf{S}_{K-1} \end{pmatrix} \quad (\text{E31})$$

Where the elements of matrices  $\mathbf{S}_u$  are defined as:

$$(S_u)_{nn'} = \pi^2 \sqrt{(2n+1)(2n'+1)} e^{-a_u - c_u - q(\mu_u - 2\pi\tau\Delta l k_0^{(u)})} \sum_{p,k,k',r} (2p+1)(2k+1)(2k'+1) e^{ir(\omega_u - \frac{\pi}{2})} \times \\ \times I_r\left(c_u \sqrt{1 + \chi_u^2}\right) I_r\left(\tau \left(1 - \frac{\lambda_u}{2\pi}\right)\right) \times \mathcal{L}_r^p(-a_u) \mathcal{L}_r^k(-b_u f) \mathcal{L}_r^{k'}(0) \times \begin{pmatrix} p & k & n \\ -r & r & 0 \end{pmatrix}^2 \begin{pmatrix} p & k' & n' \\ -r & r & 0 \end{pmatrix}^2 \quad (\text{E32})$$

Here parameters  $a_u$ ,  $b_u$ ,  $c_u$  and  $\lambda_u$  have the same meaning as before, describing the geometric and elastic properties of the DNA segments in the respective structural states ( $0 \leq u \leq K-1$ );  $\chi_u = \frac{\tau\lambda_u}{2\pi c_u}$  and  $\omega_u = \tan^{-1}(\chi_u)$ ;  $\mu_u$  is the free energy difference between the  $u^{\text{th}}$  and  $0^{\text{th}}$  structural states of a torsionally relaxed DNA calculated per single base-pair (i.e., energy price required for the DNA transition from state 0 to state  $u$ ), and  $\Delta lk_0^{(u)} = lk_{0,u} - lk_{0,0}$  is the difference between the relaxed linking numbers of DNA in states  $u$  ( $lk_{0,u}$ ) and 0 ( $lk_{0,0}$ ) calculated per single DNA base-pair.

As for the boundary condition block-matrix,  $\hat{\mathbf{V}}$ , it has the following form:

$$\hat{\mathbf{V}} = \begin{pmatrix} \mathbf{V}_0 \\ e^{-q(\mu_1 - 2\pi\tau\Delta lk_0^{(1)})} \mathbf{V}_1 \\ \vdots \\ e^{-q(\mu_u - 2\pi\tau\Delta lk_0^{(u)})} \mathbf{V}_u \\ \vdots \\ e^{-q(\mu_{K-1} - 2\pi\tau\Delta lk_0^{(K-1)})} \mathbf{V}_{K-1} \end{pmatrix} \quad (\text{E33})$$

Where in the case of the DNA ends  $\mathbf{z}_0$ -axis collinear boundary condition we have:

$$(V_u)_{nn'} = e^{b_u f} \frac{\sqrt{(2n+1)(2n'+1)}}{4\pi} \quad (\text{E34})$$

And for the DNA ends free orientation boundary condition:

$$(V_u)_{nn'} = \delta_{n'0} \times 8\pi^2 i_n(b_u f) \sqrt{2n+1} \quad (\text{E35})$$

Finally, block-matrix  $\mathbf{U}$  still has the same form as before only this time the identity matrix  $\mathbf{I}$  being repeated  $K$  times in a row:

$$\mathbf{U} = (\mathbf{I} \ \mathbf{I} \ \cdots \ \mathbf{I}) \quad (\text{E36})$$

Thus, the partition function of a mechanically stretched and twisted DNA in the general case can be calculated via the following 4 simple steps:

- 1) compute elements  $(S_u)_{nn'}$  and  $(V_u)_{nn'}$  of matrices  $\mathbf{S}_u$  and  $\mathbf{V}_u$  for  $0 \leq n, n' \leq M-1$  and  $0 \leq u \leq K-1$  by using eq. (E32) and eq. (E34) or (E35);
- 2) form square  $M \times M$  identity matrix  $\mathbf{I}$ ;
- 3) compose matrices  $\hat{\mathbf{S}}$ ,  $\hat{\mathbf{V}}$  and  $\mathbf{U}$  from the above blocks (eq. (E31), (E33) and (E36));
- 4) calculate the trace of the matrices product according to eq. (E30) to obtain the value of the DNA partition function.

## Appendix F: Vector-valued integration

In Appendix E, we have used the notion of vector-valued integration to derive eq. (E30) for the DNA partition function in the case when its segments can transit from one structural state to another. Namely, during one of the stages we have constructed the transfer matrix of DNA by utilizing the following two properties of the vector-valued integration: 1) the integration procedure can be factored out from a vector resulting in a vector-valued integral; 2) any linear functional defined on  $n$ -dimensional real space,  $\mathbb{R}^n$ , can be factored out from an integral as well, again resulting in a vector-valued integration. This section provides references to the literature, where interested readers can find useful information about the vector-valued integration technique, and gives additional insights into the properties mentioned above. For the sake of formulas simplicity, here we consider only the case of real 2-dimensional vectors and linear functionals defined on  $\mathbb{R}^2$  space. However, it should be noted that all of the equations described below can be easily extended to the general case of  $n$ -dimensional vectors and linear functionals using absolutely the same mathematical logic.

First, from the definition of the real 2D-plane ( $\mathbb{R}^2$ ) topology it immediately follows that  $\mathbb{R}^2$  is a Frechet topological vector space (i.e., it is a locally convex space, whose topology is induced by complete invariant metric [43]). Therefore, by the corollary of the Hahn-Banach separation theorem (see sec. 3.4 in ref. [43]) the dual space,  $\mathbb{R}^{2*}$ , of all of the continuous linear functionals separates points on the real 2D-plane,  $\mathbb{R}^2$ .

Second, SO(3) rotation group is a compact Lie group possessing a Borel probability measure,  $\rho$ , defined as (see sec. 51-54, 57-60 in ref. [44] and sec. I.3 in ref. [25]):

$$\rho(G) = \frac{1}{8\pi^2} \int_G d\mathbf{R} = \frac{1}{8\pi^2} \int_G \sin \beta \, d\alpha \, d\beta \, d\gamma \quad (\text{F1})$$

Where  $G$  is an arbitrary Borel subset of SO(3) group.

Finally, based on the definition of the product topology it can be shown that if  $g(\mathbf{R})$  and  $h(\mathbf{R})$ ,  $\mathbf{R} \in \text{SO}(3)$ , are any two continuous mappings of SO(3) group into the real space (i.e.,  $\text{SO}(3) \rightarrow \mathbb{R}$ ), then the following  $\text{SO}(3) \rightarrow \mathbb{R}^2$  function is also continuous:

$$\mathbf{R} \rightarrow \begin{pmatrix} g(\mathbf{R}) \\ h(\mathbf{R}) \end{pmatrix} \quad (\text{F2})$$

Using the above properties of SO(3) and  $\mathbb{R}^2$  spaces, it is then possible to correctly define the vector-valued integration on SO(3) group (see sec. 3.26 and 3.27 in ref. [43]):

$$\begin{pmatrix} x \\ y \end{pmatrix} = \int d\mathbf{R} \begin{pmatrix} g(\mathbf{R}) \\ h(\mathbf{R}) \end{pmatrix} \quad (\text{F3})$$

such that for any linear functional  $\Lambda \in \mathbb{R}^{2*}$ :

$$\Lambda \begin{pmatrix} x \\ y \end{pmatrix} = \int d\mathbf{R} \Lambda \begin{pmatrix} g(\mathbf{R}) \\ h(\mathbf{R}) \end{pmatrix} \quad (\text{F4})$$

Here  $x$  and  $y$  are real numbers ( $x, y \in \mathbb{R}$ ).

It should be noted that for any  $\Lambda \in \mathbb{R}^{2*}$  there exist a unique one-form  $(u \ v)$ , where  $u$  and  $v \in \mathbb{R}$ , such that:

$$\Lambda \begin{pmatrix} x \\ y \end{pmatrix} = (u \ v) \times \begin{pmatrix} x \\ y \end{pmatrix} = ux + vy \quad (\text{F5})$$

Substituting eq. (F5) into eq. (F4), we get:

$$\Lambda \begin{pmatrix} x \\ y \end{pmatrix} = \int d\mathbf{R} \Lambda \begin{pmatrix} g(\mathbf{R}) \\ h(\mathbf{R}) \end{pmatrix} = u \int d\mathbf{R} g(\mathbf{R}) + v \int d\mathbf{R} h(\mathbf{R}) = (u \ v) \times \begin{pmatrix} \int d\mathbf{R} g(\mathbf{R}) \\ \int d\mathbf{R} h(\mathbf{R}) \end{pmatrix} = \Lambda \begin{pmatrix} \int d\mathbf{R} g(\mathbf{R}) \\ \int d\mathbf{R} h(\mathbf{R}) \end{pmatrix} \quad (\text{F6})$$

Since eq. (F6) is valid for any  $\Lambda \in \mathbb{R}^{2*}$  and since  $\mathbb{R}^{2*}$  separates points on  $\mathbb{R}^2$  space it immediately follows that:

$$x = \int d\mathbf{R} g(\mathbf{R}) \quad \text{and} \quad y = \int d\mathbf{R} h(\mathbf{R}) \quad (\text{F7})$$

Substituting eq. (F7) into eq. (F3), it is then easy to see that:

$$\int d\mathbf{R} \begin{pmatrix} g(\mathbf{R}) \\ h(\mathbf{R}) \end{pmatrix} = \begin{pmatrix} \int d\mathbf{R} g(\mathbf{R}) \\ \int d\mathbf{R} h(\mathbf{R}) \end{pmatrix} \quad (\text{F8})$$

Although eq. (F8) is the result of sophisticated topological properties of SO(3) and  $\mathbb{R}^2$  spaces, its meaning is rather simple. Since every integral in some sense is a limit of sum and vectors are additive mathematical objects it seems intuitively clear that the integration procedure can be factored out from the vector. Here we used theorems from functional analysis to formulate and prove this statement for SO(3) group in a mathematically rigorous way.

Finally, it should be noted that while eq. (F8) was derived for the case of  $\mathbb{R}^2$  space, it can be shown that a very similar formula holds in the general case for arbitrary  $n$ -dimensional real space,  $\mathbb{R}^n$  ( $n \geq 2$ ), as well:

$$\int d\mathbf{R} \begin{pmatrix} g_1(\mathbf{R}) \\ \vdots \\ g_n(\mathbf{R}) \end{pmatrix} = \begin{pmatrix} \int d\mathbf{R} g_1(\mathbf{R}) \\ \vdots \\ \int d\mathbf{R} g_n(\mathbf{R}) \end{pmatrix} \quad (\text{F9})$$



Where  $g_1(\mathbf{R}), \dots, g_n(\mathbf{R})$  are real-valued continuous functions defined on  $\text{SO}(3)$  group.

In addition, from the definition of the vector-valued integration similarly to eq. (F4), (F5) and (F7) we have:

$$\int d\mathbf{R} (u_1 \cdots u_n) \times \begin{pmatrix} g_1(\mathbf{R}) \\ \vdots \\ g_n(\mathbf{R}) \end{pmatrix} = (u_1 \cdots u_n) \times \int d\mathbf{R} \begin{pmatrix} g_1(\mathbf{R}) \\ \vdots \\ g_n(\mathbf{R}) \end{pmatrix} \quad (\text{F10})$$

Where  $u_1, \dots, u_n$  are arbitrary real numbers.

### Appendix G: Metropolis-Monte-Carlo calculations

To generate supercoiled DNA conformations, we used the aforementioned semiflexible polymer chain model in which DNA is represented by a discretized polygonal chain consisting of  $N = 3400$  short segments of equal length,  $b = 0.5$  nm. In order to simulate the typical magnetic tweezers setup of DNA stretching / twisting experiments, the polymer chain was restricted by an impenetrable wall representing the surface of a glass coverslip at one end (parallel to the  $\mathbf{x}_0\mathbf{y}_0$ -plane) and an impenetrable bead with diameter of  $3 \mu\text{m}$  at the other end (see Figure 8).

A set of equilibrium DNA conformations was sampled using the Metropolis-Monte-Carlo procedure [45]. In these calculations, the DNA chain segments were allowed to pass through each other to generate conformations that correspond to DNA states appearing in the transfer-matrix computations.

At each iteration, the total energy of the DNA was estimated using eq. (1) and (3) as described in the main text. The DNA writhe number ( $Wr^F$ ) was calculated using the Fuller's approximation [22] by first evaluating the total DNA linking number change ( $\Delta Lk^F$ ) via eq. (4) and then subtracting the DNA twist number ( $\Delta Tw$ ) from the obtained value of  $\Delta Lk^F$  ( $Wr^F = \Delta Lk^F - \Delta Tw$ ). We also examined how much  $Wr^F$  deviates from the writhe number  $Wr = Wr^G$  estimated using a numerical approximation of the Gauss double integral, see ref. [46, 47]:

$$Wr = Wr^G = \frac{1}{4\pi} \int_0^L \int_0^L \left( \left[ \frac{d\mathbf{r}(s)}{ds} \times \frac{d\mathbf{r}(s')}{ds'} \right] \cdot \frac{\mathbf{r}(s) - \mathbf{r}(s')}{|\mathbf{r}(s) - \mathbf{r}(s')|^3} \right) ds ds' \quad (\text{G1})$$

Where  $\mathbf{r}(s)$  and  $\mathbf{r}(s')$  are the position vectors of points situating on the DNA contour that correspond to the arc lengths  $s$  and  $s'$ , respectively.

Because  $Wr^G$  is not defined for an open contour, we virtually closed the DNA chain by adding a complementing loop (see Figure 8, dashed line).  $Wr^G$  was then calculated by summing up the evaluations of the Gauss double integral over all segments composing the closed DNA chain [46]:

$$Wr^G = \frac{1}{4\pi} \sum_{j,j'=1}^N \Omega_{jj'} \quad (\text{G2})$$

Here  $j$  and  $j'$  are indexes enumerating the DNA chain segments;  $N$  is the total number of these segments, and  $\Omega_{jj'}$  is the solid angle corresponding to the quadrangle formed by segments  $j$  and  $j'$  [46].

To evaluate the contribution of each DNA subchain shown on Figure 4(b) to the writhe number ( $Wr^F$  and  $Wr^G$ ), we used the following formulas:

$$(Wr^F)_{kl} = \sum_{j=k}^l (\Delta Lk_j^F - \Delta Tw_j) \quad \text{and} \quad (Wr^G)_{kl} = \frac{1}{4\pi} \sum_{j=k}^l \sum_{j'=1}^N \Omega_{jj'} \quad (\text{G3})$$

where  $k$  and  $l$  are indices of the vertices bounding the DNA subchain;  $\Delta Tw_j(\mathbf{R}_j, \mathbf{R}_{j+1})$  and  $\Delta Lk_j^F = \frac{1}{2\pi}(\tilde{\alpha}_{j+1} + \tilde{\gamma}_{j+1} - \tilde{\alpha}_j - \tilde{\gamma}_j)$  are the local contributions to the DNA twist and linking number change by the  $j^{\text{th}}$  and  $(j+1)^{\text{th}}$  DNA segments.

In all of our calculations, the bending and twisting persistence lengths of DNA were taken to be  $A_B = 50$  nm and  $C_B = 95$  nm, the environmental temperature was set to 300 K, and  $\lambda_B = 4.3$ .

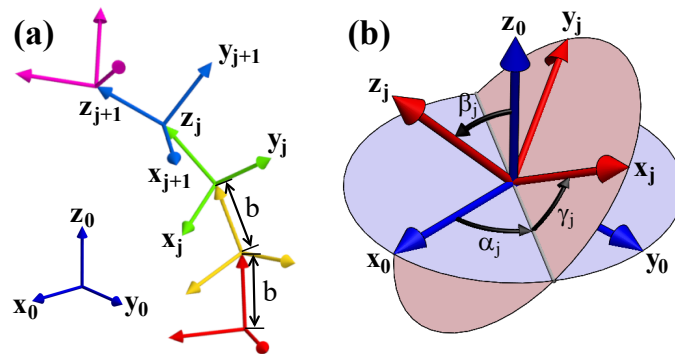


FIG. 1. **Discretized model of DNA.** (a) In the model, DNA is represented by a polygonal chain comprised of straight segments of equal length,  $b$ . Each DNA segment is considered as a rigid body with attached local Cartesian coordinate frame  $(\mathbf{x}_j, \mathbf{y}_j, \mathbf{z}_j)$ , whose 3D-orientation is described by rotation matrix  $\mathbf{R}_j$  with respect to the fixed global coordinate system  $(\mathbf{x}_0, \mathbf{y}_0, \mathbf{z}_0)$ . (b) Three Euler rotation angles  $(\alpha_j, \beta_j, \gamma_j)$  corresponding to matrix  $\mathbf{R}_j$  determine the orientation of the  $j^{\text{th}}$  DNA segment.

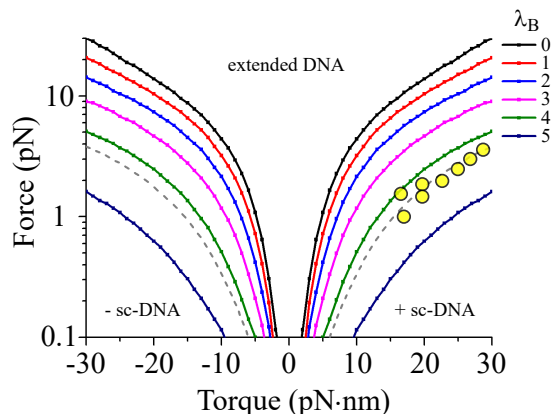


FIG. 2. **Supercoiling of B-DNA.** In the absence of B-DNA transitions into alternative structural states, the DNA phase diagram is symmetric with respect to negative and positive torques. At large forces, DNA preferentially stays in the extended conformation; whereas, at low forces it collapses into a supercoiled configuration via development of either positive (+sc-DNA) or negative (-sc-DNA) writhes, depending on the direction (i.e., sign) of the applied torque. On the plot, solid curves show the transition boundaries between the extended and supercoiled DNA states for different values of the scaling parameter  $\lambda$ , which increases from 0 to 5 from the top to the bottom curve. It can be seen from the figure that the semiflexible polymer model of DNA described in the main text accurately fits the experimentally measured transition boundary between the extended and supercoiled B-DNA states (circles), assuming that  $\lambda = 4.3$  (dashed curve). The experimental data points presented on the graph were digitized from ref. [5, 7, 9].

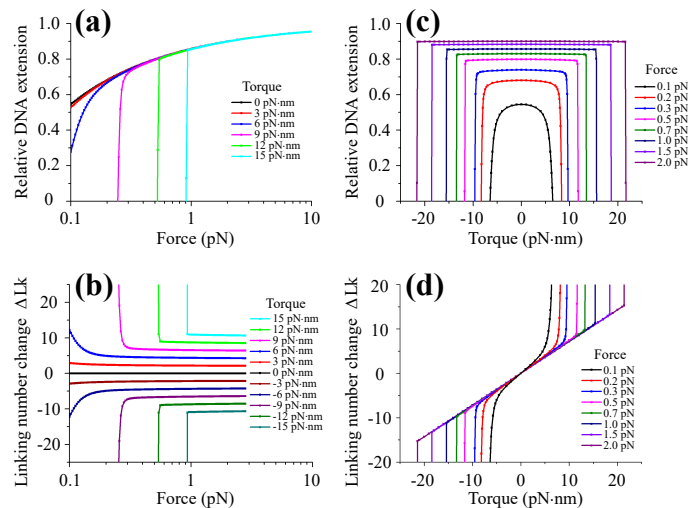


FIG. 3. **Mechanical response of B-DNA to the applied force and torque constraints.** (a, b) B-DNA force-extension curves,  $z_{\tau_0}(f)$ , and force-linking number change curves,  $\Delta Lk_{\tau_0}(f)$ , calculated at various values of the applied torque ( $\tau = \tau_0$ ). DNA transition into a supercoiled conformation is indicated by a steep DNA extension drop and simultaneous drastic change in the DNA linking number. The torque magnitude increases from the left to the right curve in panel (a) and from the bottom to the top curve in panel (b). (c, d) B-DNA torque-extension curves,  $z_{f_0}(\tau)$ , and torque-linking number change curves,  $\Delta Lk_{f_0}(\tau)$ , obtained at different values of the applied force ( $f = f_0$ ). The force magnitude increases from the bottom to the top curve in panel (c) and from the inner to the outer curve in panel (d). On panels (a) and (c), the DNA extension is normalized to the total DNA contour length.

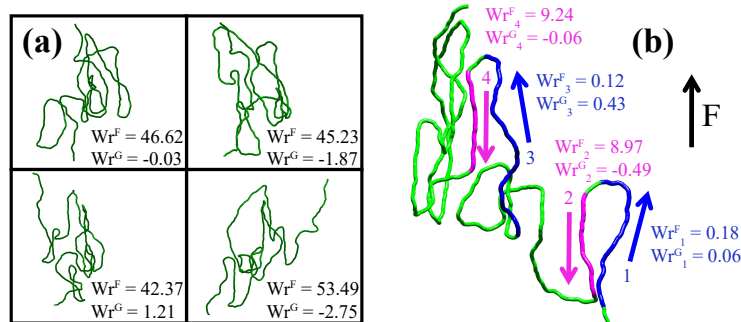


FIG. 4. **Numeric simulation of supercoiled B-DNA conformations.** (a, b) Typical configurations of B-DNA at  $f = 0.2$  pN and  $\tau = 30$  pN-nm generated using the Metropolis-Monte-Carlo algorithm [48] based on the Fuller's formula approximation. It can be seen from the figure that upon application of sufficiently large torques DNA molecule predominantly collapses into conformations in which the DNA duplex changes its orientation in alternating up-and-down manner. On panel (a),  $Wr^G$  and  $Wr^F$  are the DNA writhe numbers of the corresponding DNA polymers estimated using the double Gauss integral and the Fuller's formula, respectively. Similarly, on panel (b),  $Wr_n^G$  and  $Wr_n^F$  show the contribution of labelled DNA subchains indexed by  $n = 1-4$  to the total DNA writhe number.

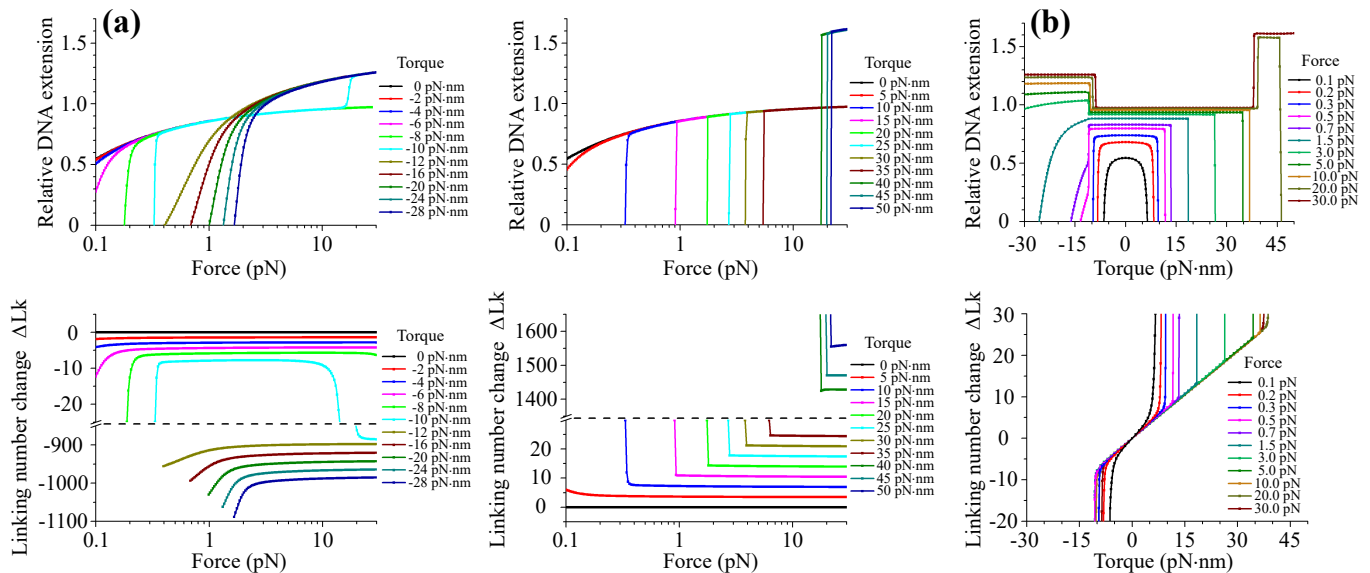


FIG. 5. **Structural stability of DNA under various mechanical constraints.** (a) DNA force-extension curves,  $z_{\tau_0}(f)$ , and force-linking number change curves,  $\Delta Lk_{\tau_0}(f)$ , calculated at different values of the applied torque ( $\tau = \tau_0$ ). Left and right panels show results obtained for negative and positive torques, respectively. The torque magnitude increases from the left to the right curve on the top panels, from the top to the bottom curve on the left-bottom panel, and from the bottom to the top curve on the right-bottom panel. Abrupt change in the DNA force-extension curves' pattern at large negative ( $\tau_0 < -10$  pN·nm) or positive ( $\tau_0 > 35$  pN·nm) torques indicated by a longer DNA extension at high forces marks the DNA transitions from B-form into alternative L- and P-DNA structures. The bottom graphs show that these transitions are accompanied by a drastic DNA linking number change due to the large difference in the relaxed linking numbers of the corresponding DNA forms. (b) DNA torque-extension curves,  $z_{f_0}(\tau)$ , and torque-linking number change curves,  $\Delta Lk_{f_0}(\tau)$ , at different values of the applied force ( $f = f_0$ ). The force magnitude increases from the bottom to the top curve on the top panel and from the inner to the outer curve on the bottom panel. At  $f_0 < 0.5$  pN all torque-extension curves have symmetric profiles with respect to both positive and negative torques, while at larger forces ( $f_0 \geq 0.5$ -0.7 pN) this symmetry breaks due to B-DNA switching into alternative L- and P-DNA structures. On panels (a) and (b), the DNA extension is normalized to the total contour length of DNA in B-form.

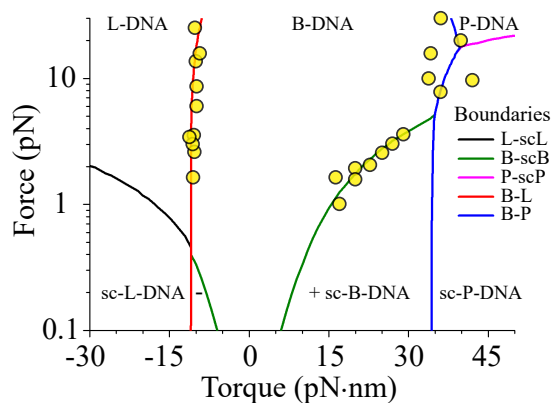


FIG. 6. **DNA phase diagram.** Solid curves indicate transition boundaries between extended (B, L and P) and supercoiled (sc-B, sc-L and sc-P) states of DNA predicted by the transfer-matrix calculations. As can be seen from the figure, the resulting phase diagram demonstrates good agreement with the existing experimental data (circles), which were digitized from ref. [4, 5, 7-9, 12, 28].

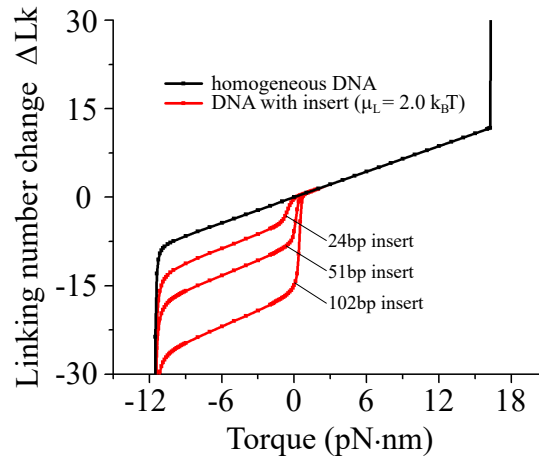


FIG. 7. **Response of a mechanically stretched heterogeneous DNA to the applied torque.** The figure demonstrates the torque-dependent change in the linking number of DNA that contains a small insert (of 24 bp, 51 bp, or 102 bp size), which is characterized by a lower energy cost of the B-to-L transition ( $\mu_L = 2.0 k_B T$ ) comparing to that of the rest of the DNA ( $\mu_L = 5.0 k_B T$ ). In the calculations, the DNA was stretched by a constant force of 1.6 pN.

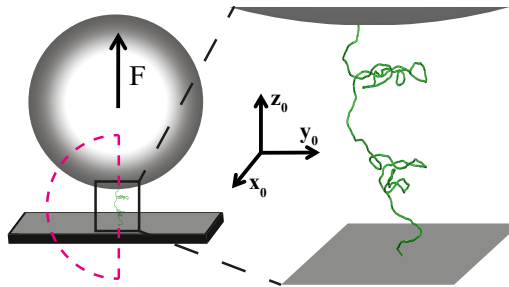


FIG. 8. **Schematic diagram of the computational setup used in the Metropolis-Monte-Carlo simulations.** In these calculations, the discretized DNA polymer chain was restricted by an impenetrable bead at one end and a wall in the  $x_0 y_0$ -plane at the other end. Stretching force,  $F$ , was applied to the bead along the  $z_0$ -axis direction. The dashed line shows a loop that was added to the DNA chain in order to form a closed contour for the writhe number ( $Wr^G$ ) evaluation.