bioRxiv preprint doi: https://doi.org/10.1101/422683; this version posted September 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Twist-bend coupling and the statistical mechanics of DNA: perturbation theory and beyond

Stefanos K. Nomidis,[1,2] Enrico Skoruppa,[1] Enrico Carlon,[1] and John F. Marko[3]

[1] *KU Leuven, Institute for Theoretical Physics, Celestijnenlaan 200D, 3001 Leuven, Belgium*
[2] *Flemish Institute for Technological Research (VITO), Boeretang 200, B-2400 Mol, Belgium*
[3] *Department of Physics and Astronomy, and Department of Molecular Biosciences, Northwestern University, Evanston, Illinois 60208, USA*

(Dated: September 19, 2018)

The simplest model of DNA mechanics describes the double helix as a continuous rod with twist and bend elasticity. Recent work has discussed the relevance of a little-studied coupling G between twisting and bending, known to arise from the groove asymmetry of the DNA double helix. Here, the effect of G on the statistical mechanics of long DNA molecules subject to applied forces and torques is investigated. We present a perturbative calculation of the effective torsional stiffness C_{eff} for small twist-bend coupling. We find that the "bare" G is "screened" by thermal fluctuations, in the sense that the low-force, long-molecule effective free energy is that of a model with $G = 0$, but with long-wavelength bending and twisting rigidities that are shifted by G-dependent amounts. Using results for torsional and bending rigidities for freely-fluctuating DNA, we show how our perturbative results can be extended to a non-perturbative regime. These results are in excellent agreement with numerical calculations for Monte Carlo "triad" and molecular dynamics "oxDNA" models, characterized by different degrees of coarse-graining, validating the perturbative and non-perturbative analyses. While our theory is in generally-good quantitative agreement with experiment, the predicted torsional stiffness does systematically deviate from experimental data, suggesting that there are as-yet-uncharacterized aspects of DNA twisting-stretching mechanics relevant to low-force, long-molecule mechanical response, which are not captured by widely-used coarse-grained models.

I. INTRODUCTION

In vivo, double-stranded DNA is typically found in a highly-deformed state, which is in part due to the interaction with the many proteins that bend and twist the double helix, but in part due to thermally-driven deformations. A substantial effort has been devoted to the study of many aspects of DNA mechanics, such as its response to applied twist and bending deformations [1]. These studies often rely on homogeneous elastic models, which, despite their simplicity, describe many aspects of single-molecule experiments [2–6], and are widely used to describe mechanical and statistical-mechanical properties of DNA (see e.g. Refs. [7–11]).

One of the simplest models describing DNA deformations is the twistable wormlike chain (TWLC), which describes the double helix as an inextensible rod, for which twist and bend deformations are independent. Symmetry arguments suggest that the TWLC is incomplete: The inherent asymmetry of the DNA molecular structure, with its major and minor grooves, gives rise to a coupling G between twisting and bending [12]. Only a limited number of studies have considered the effect of twist-bend coupling on DNA mechanics [13–16]. A systematic analysis of coarse-grained models with and without groove asymmetry have highlighted several effects associated with twist-bend coupling at long [15] and short [16] length scales. Here, we aim to clarify the role of G in the statistical mechanics of long DNA molecules, as analyzed in optical and magnetic tweezers.

We focus on analytical and numerical results for the stretching and torsional response of DNA with twist-bend coupling interaction $G \neq 0$. We first present a perturbative expansion for the partition function of the molecule, in which G is treated as the small parameter. The lowest-order results show that twist-bend coupling softens the torsional and bending stiffnesses of the double helix, recovering prior results from entirely different calculations [14]. Our new calculations reveal the existence of a previously-unidentified large force scale f_0; for forces below this scale, the bare elastic constants - including G - are not directly accessible in stretching and twisting experiments. Instead, for forces below f_0, only renormalized bending and twisting stiffnesses - which do depend on G - are observed. Because $f_0 \approx 600$ pN, the renormalized elastic model - which is the $G = 0$ TWLC - will be observed in essentially all conceivable single-molecule experiments. Thus, G is "screened", effectively renormalized to $G = 0$, in single-molecule DNA mechanics experiments.

Prior work [14] suggests a strategy to generalize our results beyond perturbation theory, to the regime where DNA is stretched by forces less than f_0. We validate both the perturbative and non-perturbative results using numerical calculations corresponding to commonly-used coarse-grained DNA elasticity models; our results turn out to closely describe results of those numerical models. Given this validation, we turn to experimental data which are reasonably well described by the low-force model, but for which there remain discrepancies, suggesting effects beyond simple harmonic elastic models like the TWLC.

bioRxiv preprint doi: https://doi.org/10.1101/422683; this version posted September 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

II. ELASTICITY MODELS OF DNA

To describe the conformation of a continuous, inextensible, twistable elastic rod, one can associate a local orthonormal frame of three unit vectors $\{\widehat{\mathbf{e}}_i\}$ $(i = 1, 2, 3)$ with every point along the rod (Fig. 1). In a continuous representation of DNA, the common convention is to choose $\widehat{\mathbf{e}}_3$ tangent to the curve and $\widehat{\mathbf{e}}_1$ pointing to the DNA groove. The frame is completed with a third vector, defined as $\widehat{\mathbf{e}}_2 = \widehat{\mathbf{e}}_3 \times \widehat{\mathbf{e}}_1$. An unstressed B-form DNA corresponds to a straight, twisted rod, with the tangent $\widehat{\mathbf{e}}_3$ being constant, and with $\widehat{\mathbf{e}}_1$ and $\widehat{\mathbf{e}}_2$ rotating uniformly about it, with a full helical turn every $l \approx 3.6$ nm, or equivalently every 10.5 base pairs.

Any deformation from this unstressed configuration can be described by a continuous set of rotation vectors $\mathbf{\Omega}$ connecting adjacent local frames $\{\widehat{\mathbf{e}}_i\}$ along the rod, using the differential equation

$$\frac{\mathrm{d}\widehat{\mathbf{e}}_i}{\mathrm{d}s} = (\mathbf{\Omega} + \omega_0 \widehat{\mathbf{e}}_3) \times \widehat{\mathbf{e}}_i, \tag{1}$$

where the internal parameter s denotes the arc-length coordinate (Fig. 1), and $\omega_0 = 2\pi/l \approx 1.75$ nm^{-1} is the intrinsic twist of DNA. Upon setting $\mathbf{\Omega} = \mathbf{0}$, one obtains the unstressed configuration mentioned above. Thus, a nonzero rotation vector $\mathbf{\Omega}(s) \neq \mathbf{0}$ corresponds to a local deformation at s around this ground state. Defining $\Omega_i \equiv \widehat{\mathbf{e}}_i \cdot \mathbf{\Omega}$, it follows that $\Omega_1(s)$ and $\Omega_2(s)$ describe local bending deformations, while $\Omega_3(s)$ describes twist deformations. In the remainder of the paper the s-dependence of $\mathbf{\Omega}$ will be implicit.

Symmetry analysis of the DNA molecule requires the energy functional E to be invariant under the transformation $\Omega_1 \to -\Omega_1$, with the consequence that [12]

$$\beta E = \frac{1}{2}\int_0^L \mathrm{d}s \left(A_1\Omega_1^2 + A_2\Omega_2^2 + C\Omega_3^2 + 2G\Omega_2\Omega_3\right), \tag{2}$$

where $\beta \equiv 1/k_\mathrm{B}T$ is the inverse temperature, A_1 and A_2 the bending stiffnesses, C the torsional stiffness and G the twist-bend coupling constant. These coefficients have dimensions of length, and can be interpreted as the contour distance along the double helix over which significant bending and twisting distortions can occur by thermal fluctuations. Our perturbative calculation will use the isotropic-bending version of this model ($A_1 = A_2 = A$), which is described by the following energy functional

$$\beta E = \frac{1}{2}\int_0^L \mathrm{d}s \left[A\left(\frac{\mathrm{d}\widehat{\mathbf{e}}_3}{\mathrm{d}s}\right)^2 + C\Omega_3^2 + 2G\Omega_2\Omega_3\right]. \tag{3}$$

Here, we have used Eq. (1) to express the sum $\Omega_1^2 + \Omega_2^2$ as the derivative of the tangent vector. The TWLC is obtained by setting $G = 0$ in Eqs. (2) and (3), corresponding to the anisotropic and isotropic cases, respectively.

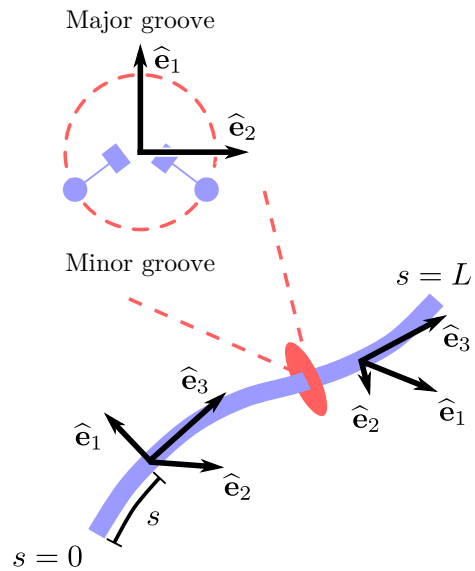

FIG. 1. Bottom: The configuration of a twistable elastic rod can be mathematically described with an orthonormal set of vectors $\{\widehat{\mathbf{e}}_1, \widehat{\mathbf{e}}_2, \widehat{\mathbf{e}}_3\}$ assigned to every point s along the rod. The vector $\widehat{\mathbf{e}}_3$ is the tangent to the curve, and describes the bending fluctuations along the rod. Top: Cross-section of the rod, indicating how the remaining vectors $\widehat{\mathbf{e}}_1$ and $\widehat{\mathbf{e}}_2$, which describe the torsional state, may be chosen in the particular case of DNA.

III. EFFECTIVE TORSIONAL STIFFNESS

In a typical magnetic tweezers experiment, a single DNA molecule of $10^3 - 10^4$ bases is attached to a solid substrate and to a paramagnetic bead at its two ends (Fig. 2). The molecule can be stretched by a linear force f and over- or undertwisted by an angle θ. The resulting torque τ exerted by the bead, which can be experimentally measured [17–20], is linear in θ for small θ

$$\tau \approx \frac{k_\mathrm{B}TC_\mathrm{eff}}{L}\theta. \tag{4}$$

Here C_eff is the *effective* torsional stiffness (in contrast to the intrinsic stiffness C), and represents the central quantity of interest here. It expresses the resistance of the DNA to a global torsional deformation, applied at its two ends.

As discussed in more detail below, C_eff is in general lower than its intrinsic equivalent C. More specifically, at low stretching forces the bending fluctuations can absorb a significant part of the applied torsional stress, leading to a globally-reduced torsional resistance $C_\mathrm{eff} < C$. On the other hand, when the applied force is sufficiently large, bending fluctuations are mostly suppressed, and hence the effective torsional stiffness tends to approach the intrinsic one. As a consequence, C_eff is going to be a monotonically-increasing function of the stretching force.

Moroz and Nelson derived an expression of C_eff for the TWLC in the limit of high forces [4, 21]. In spite of the good qualitative agreement between the theory and early

bioRxiv preprint doi: https://doi.org/10.1101/422683; this version posted September 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

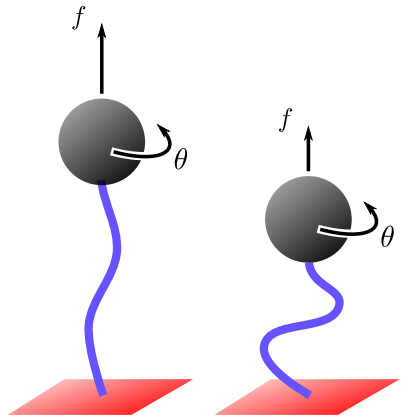

FIG. 2. Typical setup of a magnetic tweezers experiment. A DNA molecule is covalently bound to a substrate at one end and to a paramagnetic bead at the other end. An applied force f stretches the molecule, while an applied rotation θ twists it. The effective torsional stiffness $C_{\rm eff}$ is the proportionality constant connecting the applied rotation with the exerted torque [Eq. (4)].

experiments, more recent studies reported systematic deviations [14, 17, 20, 22]. For completeness we will first present in Sec. III A a short derivation of the TWLC-based theory by Moroz and Nelson. The pertubative calculation in small G is discussed in Sec. III B and generalized beyond perturbative expansion in Sec. III D.

A. The TWLC limit ($G = 0$)

In their original calculation, Moroz and Nelson [4, 21] mapped the twisted and stretched TWLC onto a quantum mechanical problem of a spinning top, and $C_{\rm eff}$ was obtained from the ground state of the associated Schrödinger equation. Here we present an alternative derivation, following the scheme illustrated in Ref. [1], which proves to be more convenient for the perturbative calculation in small G. The starting point is the partition function of a TWLC under applied force f and torque τ. The latter induces a rotation by an angle θ on the end point of the molecule (Fig. 2). The excess linking number, which we will use throughout this work, is $\Delta\mathrm{Lk} = \theta/2\pi$.

To calculate the partition function, we integrate over all possible configurations of the twistable rod, which can be parametrized by the tangent vector $\widehat{\mathbf{e}}_3(s)$ and the twist density $\Omega_3(s)$. The resulting path integral takes the form:

$$Z_0 = \int \mathcal{D}[\widehat{\mathbf{e}}_3, \Omega_3]\; e^{-\beta E_0 + \beta \mathbf{f}\cdot\mathbf{R} + 2\pi\beta\tau\Delta\mathrm{Lk}}, \tag{5}$$

where E_0 is the energy of the TWLC, obtained from Eq. (3) by setting $G = 0$, and $\mathbf{R}$ the end-to-end vector

$$\mathbf{R} = \int_0^L \widehat{\mathbf{e}}_3\, \mathrm{d}s. \tag{6}$$

We assume that the force is oriented along the z-direction, hence $\mathbf{f} = f\widehat{\mathbf{z}}$, with $\widehat{\mathbf{z}}$ a unit vector. Using a result due to Fuller [23] for open curves, the linking number can be expressed as the sum of twist and writhe, i.e. $\Delta\mathrm{Lk} = \mathrm{Tw} + \mathrm{Wr}$. The excess twist is obtained by integrating over the twist density

$$\mathrm{Tw} = \frac{1}{2\pi}\int_0^L \Omega_3(s)\,\mathrm{d}s, \tag{7}$$

while the writhe is given by [1]

$$\mathrm{Wr} = \frac{1}{2\pi}\int_0^L \frac{\widehat{\mathbf{z}}\cdot(\widehat{\mathbf{e}}_3\times \mathrm{d}\widehat{\mathbf{e}}_3/\mathrm{d}s)}{1+\widehat{\mathbf{e}}_3\cdot\widehat{\mathbf{z}}}\,\mathrm{d}s. \tag{8}$$

This representation of the writhe as a single integral (and not as a double, nonlocal integral, as in closed curves) is correct modulo 2π, and possible if the molecule is sufficiently stretched and does not loop back opposite to the direction of the applied force. Under these conditions, the denominator $1 + \widehat{\mathbf{e}}_3\cdot\widehat{\mathbf{z}}$ does not vanish, and the integral in Eq. (8) yields a finite value. Next, we insert Eqs. (6), (7) and (8) into Eq. (5), and consider the limit of strong forces and weak torques. The partition function [Eq. (5)] reduces to a Gaussian in this limit, and can be easily estimated (details can be found in Appendix A). To lowest order in τ and at large forces, one obtains the following free energy

$$F_0(f,\tau) = -fL + \sqrt{\frac{fk_{\rm B}T}{A}}L - \frac{\beta\tau^2 L}{2C_{\rm eff}} + \ldots, \tag{9}$$

where the dots denote constant or higher-order terms in τ. The effective torsional stiffness $C_{\rm eff}$ is given by (see Appendix A)

$$\frac{1}{C_{\rm eff}} = \frac{1}{C} + \frac{1}{4A}\sqrt{\frac{k_{\rm B}T}{fA}} + \ldots. \tag{10}$$

This equation was originally derived by Moroz and Nelson [4, 21]. At high forces, Eq. (10) approaches the twisted-rod limit and $C_{\rm eff} \to C$, but, in general, bending fluctuations soften the DNA torsional stiffness, so that $C_{\rm eff} < C$. The latter originates from a global coupling between torque and writhe (not to be confused with the local twist-bend coupling considered below). Note that the effect of bending fluctuations is governed by the dimensionless parameter $\sqrt{k_{\rm B}T/fA}$, which is small at low temperatures or high forces.

B. Perturbative (small-G) expansion

We now construct a perturbation expansion for the partition function, using G as the small parameter (the length scale determining whether G is "small" will be made clear below). The full partition function is

$$Z = \int \mathcal{D}[\widehat{\mathbf{e}}_3, \Omega_3]\; e^{-\beta E + \beta \mathbf{f}\cdot\mathbf{R} + \beta\tau 2\pi\Delta\mathrm{Lk}}, \tag{11}$$

bioRxiv preprint doi: https://doi.org/10.1101/422683; this version posted September 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

where now βE is given by Eq. (3) and contains a twist-bend coupling term. Assuming that G is small, we can expand the Boltzmann factor in powers of G, which gives to lowest order:

$$Z \approx Z_0 \left[1 + \frac{G^2}{2} \left\langle \left(\int_0^L \Omega_2 \Omega_3 \, \mathrm{d}s \right)^2 \right\rangle_0 + \dots \right], \quad (12)$$

where $\langle . \rangle_0$ denotes the average with respect to the unperturbed (TWLC) partition function [Eq. (5)]. Note that in the perturbative expansion, the term linear in G vanishes by the $\Omega_2 \to -\Omega_2$ symmetry of the TWLC. The full calculation of the average in the right-hand side of Eq. (12) is given in Appendix C. The final expression for the free energy is of the form [Eq. (C38)]

$$F(f,\tau) = -fL + \sqrt{\frac{f k_\mathrm{B} T}{A^*}} L + \Gamma \tau L - \frac{\beta \tau^2 L}{2 C_\mathrm{eff}} + \dots, \quad (13)$$

where terms of negligible contribution were omitted (see Appendix C).

In the above, we have introduced the rescaled bending stiffness

$$\frac{1}{A^*} = \frac{1}{A} \left(1 + \frac{G^2}{2AC} \right), \quad (14)$$

together with the parameter

$$\Gamma = \frac{G^2 d^2(f)}{4 A^2 C^2 \omega_0}. \quad (15)$$

The dimensionless, force-dependent scale factor $d(f)$ will be discussed below; we note that it appears in the coefficient Γ [Eq. (15)], but not in A^* [Eq. (14)]. Finally, the form of the effective torsional stiffness C_eff is

$$\frac{1}{C_\mathrm{eff}} = \frac{1}{C} \left[1 + \frac{G^2}{AC} d(f) \right] + \frac{1}{4A} \left(1 + \frac{3}{4} \frac{G^2}{AC} \right) \sqrt{\frac{k_\mathrm{B} T}{f A}}. \quad (16)$$

The scale factor $d(f)$ is also present in C_eff.

Examination of these formulae indicate that the expansion is in powers of the dimensionless parameter G^2/AC, which, given our current estimates for the stiffnesses ($A \approx 50$ nm, $C \approx 100$ nm, $G \approx 30$ nm), is less than 1, although we note that $G^2/AC < 1$ is a stability requirement for the microscopic energy [12]. Our computation neglects terms beyond first order in G^2/AC.

C. Effective torsional stiffness C_eff

Equation (16) is the central result of this paper, and extends the TWLC result by Moroz and Nelson [Eq. (10)], which is recovered in the the limit $G \to 0$. The perturbative corrections are governed by the dimensionless parameter G^2/AC, and give rise to a further torsional softening of the molecule, i.e. $C_\mathrm{eff}(G \neq 0) < C_\mathrm{eff}(G = 0)$, as pointed out in Ref. [14]. Equation (16) contains also a force-dependent, crossover function, which can be approximated as (see Appendix C)

$$d(f) \approx \frac{1}{1 + f/f_0}, \quad (17)$$

where $f_0 = A \omega_0^2 k_\mathrm{B} T$ is the characteristic force above which $d(f)$ starts to significantly drop below its low-force limit of $d(0) = 1$. To understand this force scale, which has no counterpart in the Moroz and Nelson formula [Eq. (10)], we recall that the correlation length for a stretched wormlike chain is $\xi = \sqrt{A k_\mathrm{B} T / f}$ [1]. Therefore, f_0 is the force associated with a correlation length of the order of the distance between neighboring bases, i.e. $\xi = 1/\omega_0$.

For DNA ($A \approx 50$ nm, $\omega_0 = 1.75$ nm^{-1}, $k_\mathrm{B} T = 4$ pN·nm) we see $f_0 \approx 600$ pN, which is far above the force where the double helix starts to be itself stretched (≈ 20 pN), force-denatured (≈ 60 pN), and is in fact comparable to where the covalently-bonded backbones will break. Hence, for forces relevant to experiments we are concerned with ($f < 10$ pN), one may simply set $d(f) \approx 1$. We will refer to this limit as the "low-force limit", but one should keep in mind that our perturbative theory is computed for the "well-stretched" limit, i.e. $f > k_\mathrm{B} T / A \approx 0.1$ pN. Therefore our perturbative theory is applicable in the force range of roughly 0.1 to 10 pN.

We emphasize that Eq. (16) can be written as

$$\frac{1}{C_\mathrm{eff}} = \frac{1}{C^*} + \frac{1}{4A^*} \sqrt{\frac{k_\mathrm{B} T}{f A^*}}, \quad (18)$$

where A^* is defined in (14), and

$$\frac{1}{C^*} \equiv \frac{1}{C} \left[1 + \frac{G^2}{AC} d(f) \right] + \mathcal{O}(G^2/AC). \quad (19)$$

Equation (18) has exactly the same form as the Moroz and Nelson formula [Eq. (10)], with rescaled bending and torsional stiffnesses. The importance of this result is paramount: since $d(f) = 1$ in the range of experimentally relevant forces, the torsional stiffness (and in fact the partition function itself) depends only on the "renormalized" stiffnesses A^* and C^*, meaning that G by itself cannot be determined from fitting of $C_\mathrm{eff}(f)$ (or any other equilibrium quantity versus f); only the effective stiffnesses A^* and C^* can be determined from experiments in the low-force regime.

D. Non-perturbative result for C_eff valid for $f < f_0$

Equation (16) has been derived on the basis of a systematic perturbation expansion in G (more formally, in the small parameter $G^2/AC < 1$). When cast in the form of Eq. (18), it is apparent that there is a simple way to extend the results to a more general, nonperturbative

bioRxiv preprint doi: https://doi.org/10.1101/422683; this version posted September 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

case, where G may be large and the bending possibly anisotropic, i.e. $A_1 \neq A_2$ in Eq. (2). The key physical idea here is that for forces below the gigantic force scale f_0, thermal fluctuations at the helix scale where G correlates bending and twisting fluctuations is unperturbed ($d(f) = 1$), and therefore we might as well just consider DNA to have the effective twisting and bending stiffnesses that is has at *zero* force.

In absence of applied torques and forces ($f, \tau = 0$), the partition function of Eq. (11) can be evaluated exactly [14]. Due to twist-bend coupling, the bending and torsional stiffnesses are renormalized as [14]:

$$\kappa_{\rm b} = A\,\frac{1-\dfrac{\varepsilon^2}{A^2}-\dfrac{G^2}{AC}\left(1+\dfrac{\varepsilon}{A}\right)}{1-\dfrac{G^2}{2AC}}, \tag{20}$$

$$\kappa_{\rm t} = C\,\frac{1-\dfrac{\varepsilon}{A}-\dfrac{G^2}{AC}}{1-\dfrac{\varepsilon}{A}}, \tag{21}$$

where we have introduced the parameters $A = (A_1 + A_2)/2$ and $\varepsilon = (A_1 - A_2)/2$. Eqs. (20) and (21) quantify the energetic cost of bending and twisting deformations, respectively, in the same way A and C do within the TWLC. Note that, by setting $G = 0$ in these expressions, one recovers the TWLC limit, $\kappa_{\rm t} = C$ and $\kappa_{\rm b} = 2A_1A_2/(A_1 + A_2)$ i.e. the renormalized bending stiffness is the harmonic mean of A_1 and A_2, which is a known result (see e.g. Refs. [24, 25]). If $G \neq 0$ one has $\kappa_{\rm b} < 2A_1A_2/(A_1 + A_2)$ and $\kappa_{\rm t} < C$, i.e. twist-bend coupling "softens" the bending and twist deformations of the DNA molecule, already if $f = 0$, $\tau = 0$. Eq. (16) then describes two different effects: one is the thermally-induced torsional softening due to bending fluctuations [already present in the TWLC expression (10)] and the other is the G-induced softening, which is captured by the two factors between parentheses in Eq. (16).

Setting $\varepsilon = 0$ and expanding Eqs. (20) and (21), one finds $\kappa_{\rm b} = A^* + \mathcal{O}(G^4)$ and $\kappa_{\rm t} = C^* + \mathcal{O}(G^4)$, which suggests the following, more general, nonperturbative result for $C_{\rm eff}$, valid as long as $f \ll f_0$

$$\frac{1}{C_{\rm eff}} = \frac{1}{\kappa_{\rm t}} + \frac{1}{4\kappa_{\rm b}}\sqrt{\frac{k_{\rm B}T}{f\kappa_{\rm b}}}. \tag{22}$$

This relation, similar to Eq. (18), has the same form as the Moroz and Nelson formula [Eq. (10)], with A and C replaced by $\kappa_{\rm b}$ and $\kappa_{\rm t}$ [much as our result for $C_{\rm eff}$ of Eq. (19) has the Moroz-Nelson form with $A \to A^*$ and $C \to C^*$]. As we will show in the next Section, this new, nonperturbative result for the continuum model is in excellent agreement with numerical Monte Carlo (MC) and molecular dynamics (MD) calculations.

E. Twist-bend-coupling-induced DNA unwinding

An intriguing feature of the perturbative calculation is the appearance of a term linear in τ in the free energy [Eq. (13)], which induces an unwinding of the helix at zero torque. In particular, from Eqs. (11), (13) and (15) it follows that

$$\langle \Delta{\rm Lk}\rangle|_{\tau=0} = -\frac{1}{2\pi}\,\frac{\partial F}{\partial \tau}\bigg|_{\tau=0} = -\frac{d^2(f)}{8\pi\omega_0}\,\frac{G^2 L}{A^2C^2}. \tag{23}$$

The scale for this thermal unwinding is very small: using typical values of DNA parameters ($A \approx 50$ nm, $C \approx 100$ nm, $G = 30$ nm, $\omega_0 = 1.75$ nm^{-1}) we find an unwinding angle per contour length of $2\pi\,\langle \Delta{\rm Lk}\rangle / L = -G^2/4\omega_0 A^2C^2 \approx -5\times10^{-6}$ rad/nm (about -1×10^{-5} degrees per base pair). This G-generated shift in helix twisting is inconsequentially small, but it is worth noting that this term is present in the perturbation theory.

It has been long known that there is a gradual unwinding of the double helix as the temperature is increased [26, 27], and this effect has been recently observed at the single-DNA level [28]. Although one might imagine an overall T^2 dependence of this term (from the factors of β in the Boltzmann factor), this dependence can only generate a tiny fraction of the observed temperature-dependent unwinding of $\approx -1\times10^{-2}$ degrees/K·bp. The experimentally observed unwinding is likely due to temperature-dependence DNA conformational changes [28], and is beyond the scope of being captured by the simple elastic models discussed here; in particular the observed unwinding of DNA with increasing temperature is not attributable to the twist-bend coupling G.

F. "Janus strip" limit ($\omega_0 \to 0$)

Equation (22) is not generally valid for any arbitrary polymer with twist-bend coupling, but its validity is linked to the physical parameters characterizing DNA elasticity. These conspire to set forces encountered in typical experiments (below 10 pN) to be far below the characteristic force $f_0 = k_{\rm B}TA\omega_0^2 \approx 600$ pN at which one starts to see effects at the helix repeat scale, i.e. force-driven unwinding of the double helix due to quenching of thermal fluctuations and the influence of G. In this sense, ω_0 can be regarded as a "large" parameter: combinations of it and the elastic constants give dimensionless constants large compared to unity, e.g. $A\omega_0$, $C\omega_0 \approx 10^2 \gg 1$ and $\sqrt{Ak_{\rm B}T/f}\,\omega_0 \gg 1$ for $f < 10$ pN).

For this reason, several ω_0-dependent terms, which in principle would contribute to $C_{\rm eff}$ at order G^2, can in practice be neglected in the application of the theory to DNA [see e.g. Eqs. (C30) and (C31) in Appendix]. The neglect of these terms leads to $C_{\rm eff}$ taking the simple form given by Eq. (18), in which A^* and C^* are the renormalized stiffnesses.

While not relevant to DNA, we might imagine other polymer structures for which ω_0 is not so large, i.e.

bioRxiv preprint doi: https://doi.org/10.1101/422683; this version posted September 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

where ω_0 is closer in size to $1/A$ or $1/C$. In this case one cannot ignore these additional terms, and $d(f) \approx 1/(1+f/f_0)$ might drop significantly over experimentally-relevant force ranges. Chiral proteins, lipid filaments, or even nanofabricated objects might comprise realizations of such situations.

As an example we consider the extreme limit $\omega_0 \to 0$, corresponding to a "Janus strip", an elastic strip with inequivalent faces (i.e. inequivalent major and minor "grooves"), and, thus, nonzero G. In this case, using the more complete and complicated results for the perturbative expansion given in the Appendix, we obtain to lowest order in G

$$\frac{1}{C_{\rm eff}} = \frac{1}{C} + \frac{1}{4A}\left[1 + \left(\frac{3}{4} + \frac{2A^2}{C^2}\right)\frac{G^2}{AC}\right]\sqrt{\frac{k_{\rm B}T}{fA}}. \quad (24)$$

In this limit, compared to the large-ω_0 case relevant to DNA, there is a more gradual shift of $C_{\rm eff}$ up to its high-force limit, and an inequivalence of the form of $C_{\rm eff}$ to the Moroz-Nelson form. Physically, this is because the intrinsic chirality of the filament is now gone, eliminating the "screening" of effects of G at low forces, and the simple dependence of the low-force themodynamics on only the coarse-grained stiffnesses $\kappa_{\rm t}$ and $\kappa_{\rm b}$. Experiments on such Janus strips, or on "soft-helix" objects where $A\omega_0$, $C\omega_0 < 1$ could provide realizations of this limit of the theory. Ref. [14] showed that Eq. (24) fits experimental data for DNA surprisingly well, despite not taking account of double helix chirality.

IV. NUMERICAL CALCULATIONS

To check the validity of the analytical results presented above, we performed numerical simulations of two different models. The first model, referred to as the *triad model*, is obtained from the discretization of the continuum elastic energy (2) and treated using MC computations. The second model is oxDNA, a coarse-grained model of nucleic acids [29], treated using MD calculations.

A. Triad model

The triad model is comprised of a series of N orthonormal vectors $\{\widehat{\mathbf{e}}_i(k)\}$ with $i = 1, 2, 3$ and $k = 0, 1, 2 \ldots N$, each representing a single base pair, interacting with its neighbors according to Eq. (2). The total length of the molecule is $L = Na$, with $a = 0.34$ nm the base pair distance. The ground state of this model is a twisted, straight rod, with $\widehat{\mathbf{e}}_3$ being aligned with the direction of the stretching force, and the vectors $\widehat{\mathbf{e}}_1$, $\widehat{\mathbf{e}}_2$ rotating about $\widehat{\mathbf{e}}_3$ with an angular frequency ω_0. A cluster move consisted of a rotation of the whole subsystem beyond a randomly-selected triad by a random angle. The new rotation vector $\mathbf{\Omega}$ was calculated based on Rodrigues' rotation formula (see e.g. Supplementary Material of Ref. [15]), then the energy was updated from a discretized version of Eq. (2), with the addition of a force term [see Eq. (5)]. The move was accepted or rejected according to the Metropolis algorithm. The stiffness constants A_1, A_2, C and G are input parameters for the model, and may, therefore, be arbitrarily chosen, provided the stability condition $G^2 < A_2C$ is met [for which the quadratic form in Eq. (2) is positive definite]

The effective torsional stiffness was calculated at zero torque from linking number fluctuations:

$$C_{\rm eff} = \frac{L}{4\pi^2\langle(\Delta{\rm Lk} - \langle\Delta{\rm Lk}\rangle)^2\rangle}, \quad (25)$$

The variance of linking number in the denominator was evaluated from the topological relation $\Delta{\rm Lk} = \Delta{\rm Tw} + {\rm Wr}$, with twist and writhe obtained from the discretization of Eqs. (7) and (8), respectively. To check the validity of our results, the writhe was also evaluated from the double-integral formula, following the method of Ref. [30], and no significant differences were found for forces > 0.25 pN. In all simulations, the size of the system was 600 triads (base pairs), above which the results remained identical within that force range.

1. *Isotropic bending*

Figure 3 shows the results of Monte Carlo calculations for the isotropic model of Eq. (3), with $A = 50$ nm, $C = 100$ nm and $G = 0, 20, 40$ nm (top to bottom). The data are plotted as a function of the dimensionless parameter $\sqrt{k_{\rm B}T/fA}$. The numerical errors are smaller than the symbol sizes, and hence not shown.

In absence of twist-bend coupling ($G = 0$, upper panel), the Monte Carlo data are in excellent agreement with the Moroz-Nelson theory. We compare them both to Eq. (10) (dashed line) and the following expression (solid line)

$$C_{\rm eff} = C\left(1 - \frac{C}{4A}\sqrt{\frac{k_{\rm B}T}{fA}}\right). \quad (26)$$

The latter is obtained from the lowest-order expansion of Eq. (10) in $\sqrt{k_{\rm B}T/fA}$, and is a straight line when plotted as a function of the rescaled variable of Fig. 3. Eqs. (10) and (26) coincide to leading order in $1/\sqrt{f}$, and any differences in the two expressions only occur at low force scales, where higher-order corrections become relevant. Eq. (26) fits the Monte Carlo data extremely well over the whole range of forces analyzed ($f \geq 0.25$ pN), while Eq. (10) deviates at low forces.

The middle panel of Fig. 3 shows Monte Carlo results for $G = 20$ nm (points), which we compare both to the results for the perturbative expansion for $1/C_{\rm eff}$, Eq. (18) (dashed line) and the following similar expansion result

bioRxiv preprint doi: https://doi.org/10.1101/422683; this version posted September 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

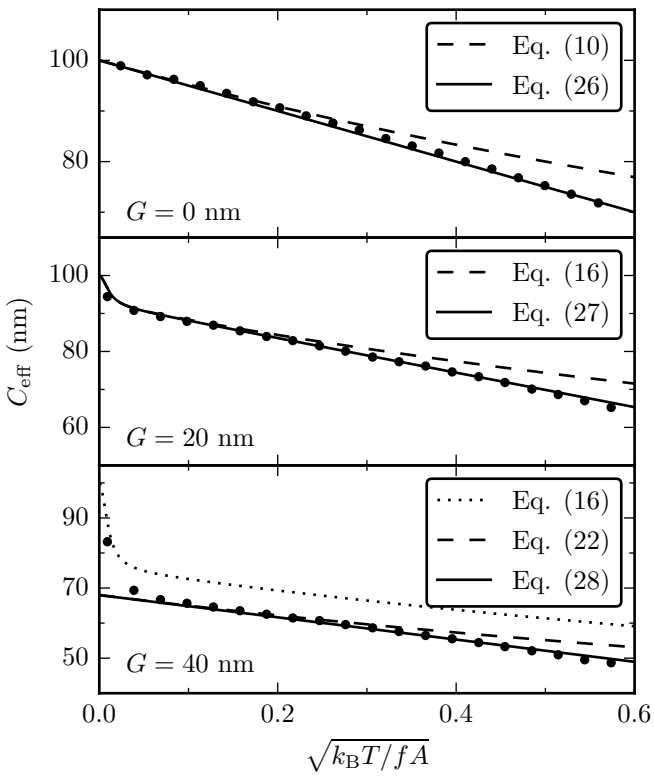

FIG. 3. Comparison of Monte Carlo simulations of the triad model (points) with various analytical expressions of $C_{\rm eff}$ (lines) for $A = 50$ nm, $C = 100$ nm and $G = 0, 20, 40$ nm. The data are plotted as a function of $\sqrt{k_{\rm B}T/fA}$ and correspond to $f \geq 0.25$ pN. The numerical results are in excellent agreement with the analytical expressions, both in the perturbative and nonperturbative regimes (see text). Error bars of Monte Carlo data are smaller than symbol sizes.

for $C_{\rm eff}$ (solid line)

$$C_{\rm eff} = \frac{C}{1 + \frac{G^2}{AC} d(f)} \left(1 - \frac{C}{4A} \frac{1 + \frac{3G^2}{4AC}}{1 + \frac{G^2}{AC} d(f)} \sqrt{\frac{k_{\rm B}T}{fA}} \right), \tag{27}$$

obtained by expanding Eq. (16) to lowest order in $1/f$. The latter is in excellent agreement with Monte Carlo data in the whole range of forces considered, indicating that $G = 20$ nm falls within the range of validity of the perturbative calculation ($G^2/AC = 0.08$ in this case). We note that the two perturbative expansion results converge together at high forces ($[k_BT/(Af)]^{1/2} \to 0$ and also show the upturn at the very highest forces associated with the force-dependence of $d(f)$.

Finally, the lower panel of Fig. 3 shows the results of Monte Carlo simulations for $G = 40$ nm. The numerical data deviate substantially from Eq. (16) (dotted line), indicating that G is leaving the range of validity of the perturbative calculation ($G^2/AC = 0.32$). The remaining curves show the nonperturbative result for $1/C_{\rm eff}$ Eq. (22) (dashed line), together with the following nonperturbative expression for $C_{\rm eff}$ (solid line)

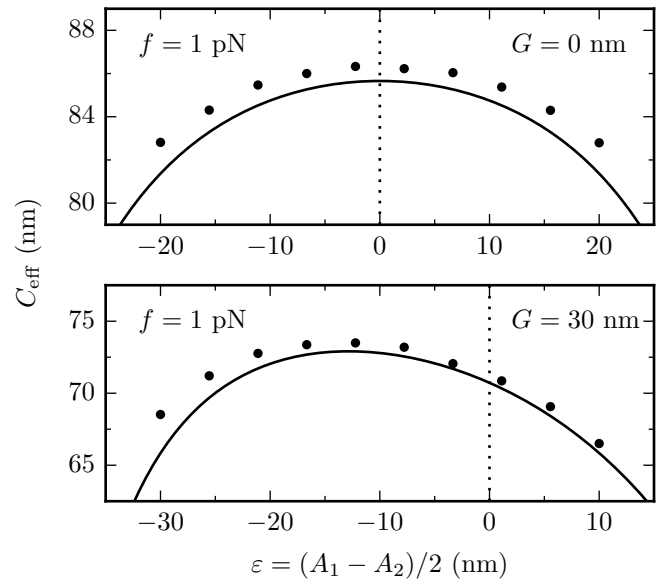

FIG. 4. Effect of bending anisotropy ($A_1 \neq A_2$) on the effective torsional stiffness at a fixed force $f = 1$ pN. Numerical data from Monte Carlo simulations of the triad model (points) are in good agreement with the analytical, nonperturbative predictions of Eq. (28) (solid lines). The vertical dashed lines indicate the isotropic case ($\varepsilon = 0$). A nonvanishing twist-bend coupling (lower panel) induces a $\varepsilon \to -\varepsilon$ symmetry breaking. Error bars of Monte Carlo data are smaller than symbol sizes.

$$C_{\rm eff} = \kappa_{\rm t} \left(1 - \frac{\kappa_{\rm t}}{4\kappa_{\rm b}} \sqrt{\frac{k_{\rm B}T}{f\kappa_{\rm b}}} \right), \tag{28}$$

the latter being in excellent agreement with Monte Carlo data for $f < f_0$, validating the nonperturbative result (note that $[k_BT/(f_0A)]^{1/2} \approx 0.01$). For $f > f_0$, the upturn of $C_{\rm eff}$ towards the bare value of C is apparent; this effect, while not given by the nonperturbative results, is present in the perturbation expansion results. We conclude that our non-perturbative result indeed provides a quantitative account of $C_{\rm eff}$ for $f < f_0$, where we expect it to be valid.

2. *Anisotropic bending*

While the perturbative calculation [Eq. (16)] was restricted to the isotropic case ($A_1 = A_2$), the nonperturbative result [Eq. (28)] has a broader range of applicability, and is able to describe the anisotropic case as well. Figure 4 shows the results of Monte Carlo calculations of $C_{\rm eff}$ for various values of the anisotropy parameter $\varepsilon = (A_1 - A_2)/2$ and a fixed value of the force $f = 1$ pN. The Monte Carlo data are in very good agreement with Eq. (28), plotted with solid lines. The differences are within 5%, and are probably due to higher order corrections in $1/f$ (recall that all analytical results are based on

bioRxiv preprint doi: https://doi.org/10.1101/422683; this version posted September 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

	A_1	A_2	C	G	κ_b	κ_t
oxDNA1	84(14)	29(2)	118(1)	<0.3	43	118
oxDNA2	85(10)	35(2)	109(1)	25(1)	44	92

TABLE I. Values of the stiffness coefficients for oxDNA1 and oxDNA2 (expressed in nm), derived from MD data of Ref. [15]. oxDNA1, which has symmetric grooves, is characterized by a negligible twist-bend coupling constant, while $G = 25$ nm for oxDNA2, which has asymmetric grooves. The values of κ_b and κ_t are obtained from Eqs. (20) and (21), respectively, and were found to agree with direct computations of those quantities [15]. Note that the oxDNA2 stiffness coefficients have been transformed to coordinates compatible with this paper, see Appendix D).

a large-force expansion). In absence of twist-bend coupling ($G = 0$), Eq. (28) is symmetric in ε, as in this case Eqs. (20) and (21) give $\kappa_b = (A^2 - \varepsilon^2)/A$ and $\kappa_t = C$, respectively. A nonzero G induces nonvanishing terms, which are linear in ε both in κ_b and κ_t [Eqs. (20) and (21)], leading to a breaking of the $\varepsilon \to -\varepsilon$ symmetry.

B. oxDNA

oxDNA is a coarse-grained model describing DNA as two intertwined strings of rigid nucleotides [29]. It has been used for the study of a variety of DNA properties, ranging from single molecules to large-scale complexes [15, 16, 29, 31–33]. To date, two versions of oxDNA exist: one with symmetric grooves (oxDNA1) [29] and one with asymmetric grooves (oxDNA2) [32]. Comparing the torsional response of the two versions will allow us to infer the effect of the groove asymmetry on $C_{\rm eff}$. Differently from the triad model, in which the stiffness constants A_1, A_2, C and G are input parameters, in oxDNA they are determined by the molecular force fields used. These force fields were accurately tuned so that the experimental DNA structural, mechanical and thermodynamic properties (as persistence length, melting temperatures and torque-induced supercoiling) are well reproduced [29]. As for real DNA, for oxDNA the elastic constants are emergent via coarse-graining of fluctuations of smaller-scale, molecular motion degrees of freedom.

The stiffness parameters of oxDNA were recently estimated from the analysis of the equilibrium fluctuations of an unconstrained molecule [15], and are shown in Table I (the values of the elastic constants for oxDNA2 shown are the result of transformation of the values obtained in Ref. [15] for the helical coordinate system used in that paper, to the non-helical coordinate system of this paper; see Appendix D). In line with the symmetry arguments of Ref. [12], twist-bend coupling is absent in oxDNA1 (symmetric grooves), while its magnitude is comparable to that of the elastic constants A_1, A_2 and C in oxDNA2 (asymmetric grooves). Table I also reports the values of κ_b and κ_t, which can be obtained in two different, yet consistent, ways [15]: either indirectly from Eqs. (20) and (21), by plugging in A_1, A_2, C and G of Table I, or directly from the analysis of the corresponding correlation functions in simulations (κ_b and $2\kappa_t$ are, respecively, the bending and twist stiffnesses [14]).

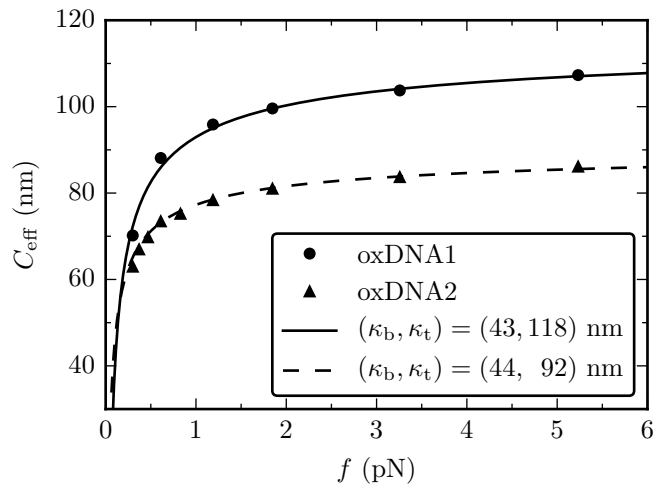

FIG. 5. Comparison of oxDNA simulations to Eq. (28). Solid and dashed lines show the nonperturbative result for $C_{\rm eff}$ for oxDNA1 and oxDNA2 values of Table I, respectively. Error bars for the oxDNA data are smaller than symbol sizes.

Figure 5 shows a plot of the effective torsional stiffness as a function of the applied force, both for oxDNA1 (circles) and oxDNA2 (triangles). $C_{\rm eff}$ was evaluated using twist fluctuations via Eq. (25). At large forces, and in agreement with the experimental evidence, oxDNA undergoes a structural transition, hence the simulations were restricted to $f \leq 10$ pN. The solid and dotted lines of Fig. 5 are plots of Eq. (28) using κ_b and κ_t from Table I. For oxDNA1 there is an excellent agreement between the nonperturbative theory and simulations. In this case the nonperturbative theory reduces to the Moroz-Nelson result, with $\kappa_t = C$ and $\kappa_b = A(1-\epsilon^2/A^2)$; the good account of oxDNA1 $C_{\rm eff}$ by this formula was noted previously (see Ref. [34], Fig. S7). We note that in the light of our present results, this good agreement validates the use of the values of the stiffness parameters obtained in Ref. [15].

V. DISCUSSION

We have investigated the effect of the twist-bend coupling G on the statistical-mechanical properties of a stretched DNA molecule, using analytical and numerical methods. Our major analytical results are based on a perturbative calculation of the effective torsional stiffness $C_{\rm eff}$, the torsional resistance of a long DNA molecule stretched by an applied force f. The calculation is valid for small values of G, and generalizes the expression derived by Moroz and Nelson, which was derived for $G = 0$ [4].

bioRxiv preprint doi: https://doi.org/10.1101/422683; this version posted September 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

A. Screening effect for $f < f_0$

A striking feature of our theory is the appearance of a large force scale, $f_0 = k_B T A \omega_0^2 \approx 600$ pN. For forces well below this gargantuan force level (essentially all single-DNA mechanical experiments concern forces far below this value) the effect of G becomes solely renormalization of the bending and twisting stiffnesses κ_b and κ_t; direct effects of twist-bend coupling are "screened" at lower force scales. Only at forces $f \gg f_0$ do the bare elastic constants start to reveal themselves: in this regime C_eff finally approaches its intrinsic value C.

This "screening" feature of the perturbative theory suggested to us that we could consider DNA for $f < f_0$ to be described by a TWLC with persistence lengths set to the zero-force long-molecule stiffnesses κ_b and κ_t. Combining formulae for the stiffnesses for freely fluctuating DNA [14] with the C_eff formula of Moroz and Nelson [4] gave us a nonperturbative formula for C_eff in terms of the elastic constants A_1, A_2, C and G. MC calculations for the triad model which discretizes the continuum elasticity theory (3) were found to be in excellent agreement both with the perturbative (i.e., small G) and nonperturbative (for larger G) expressions of C_eff. We note that despite being inaccessible experimentally, we were able to see the very high force behavior of the perturbative theory - namely the increase of C_eff from its low-force Moroz-Nelson behavior, towards its "naked" value of C in the triad MC calculations.

B. oxDNA under moderate forces is described by the TWLC plus twist-bend coupling

To test whether our analytical results describe the coarse-grained behavior of a more realistic molecular model of DNA, we carried out MD simulations of oxDNA, a coarse-grained model describing DNA as two intertwined strings of rigid nucleotides [29]. C_eff for oxDNA1, a DNA model with symmetric grooves, was determined in previous work [34] and found to be in agreement with the ($G = 0$) Moroz-Nelson theory.

For the more realistic oxDNA2, which has the asymmetric grooves of real DNA and hence twist-bend coupling ($G \neq 0$) [15], we found that C_eff is in excellent agreement with the nonperturbative theory Eq. (28) without any adjustable parameters, as the elastic constants were determined in previous work [15]. oxDNA2 appears to be precisely described by our nonperturbative theory for forces $f \ll f_0$ (recall that oxDNA undergoes internal structural transitions for forces of a few tens of pN, providing a more stringent contraint on force than the giant force scale f_0). Put another way, the "TWLC plus G" is the "correct" low-force, long-fluctuation wavelength model for oxDNA2.

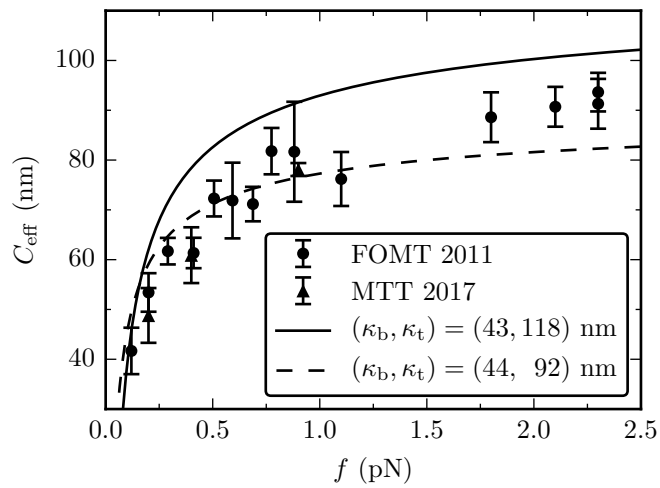

FIG. 6. Comparison of the theory from Eq. (28) (lines) with Magnetic Tweezer experiments (symbols) for C_eff vs. force. The lines have the same parametrization of the dashed and solid lines of Fig. 5, which fit extremely well oxDNA and oxDNA2 data. Two sets of experiments are shown: the freely-orbiting magnetic tweezers [22] (circles) and magnetic torque tweezers [14] (triangles).

C. Experimental data

We finally compare the analytical results with experimental magnetic tweezers data of Refs. [14, 22]. Figure 6 shows experimental data (symbols) together with plots of Eq. (28) for two sets of parameters κ_b and κ_t (lines). The latter are identical to the solid and dashed lines of Fig. 5, which are numerically precise descriptions of oxDNA1 and oxDNA2, respectively. Since the force fields in oxDNA were carefully tuned to reproduce several mechanical and thermodynamic properties of DNA [29], it is sensible to directly compare our nonperturbative theory to experimental data (Fig. 6).

As reported in previous papers [14, 17] the experimental C_eff data are systematically lower than the prediction of the Moroz and Nelson theory, which precisely matches the oxDNA1 results (solid line in Fig. 6). The oxDNA2/nonperturbative theory (dashed curve) is closer to the experimental data, especially in the low force regime $f < 1$ pN. However, some systematic deviations are noticeable at higher forces, where theory appears to underestimate the experimental C_eff. In addition, measurements at $f = 15$ pN (albeit for a slightly different assay) yield $C_\mathrm{eff} = 110$ nm [35], well above the oxDNA2 value of $C_\mathrm{eff} = 92$ nm.

We conclude by noting that oxDNA2 - which has the realistic features of groove asymmetry and $G \neq 0$ - produces data in reasonable agreement with experiments. We have also shown that in the force range where we expect that coarse-graining of oxDNA2 should agree with our analytical results, it does. In that same force range ($f < 10$ pN), oxDNA2 and our analytical results show some systematic differences that suggest that physics be-

bioRxiv preprint doi: https://doi.org/10.1101/422683; this version posted September 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

yond simple harmonic elasticity may be in play at intermediate forces (1 to 10 pN), generating torsional stiffening of DNA. This new ingredient might be sub-helix-scale structural transitions as discussed in [36], and highlighted in a recent study of thermally-driven DNA unwinding [28]. The next generation of coarse-grained DNA models likely will have to consider this kind of additional, internal degree of freedom to properly describe severe distortion of DNA by proteins, or similar situations where strong forces are applied at short length scales.

ACKNOWLEDGMENTS

Discussions with M. Laleman and T. Sakaue are gratefully acknowledged. SN acknowledges financial support from the Research Funds Flanders (FWO Vlaanderen) grant VITO-FWO 11.59.71.7N, and ES from KU Leuven Grant No. IDO/12/08. JFM is grateful to the Francqui Foundation (Belgium) for financial support, and to the US NIH through Grants R01-GM105847, U54-CA193419 and U54-DK107980.

Appendix A: TWLC at strong stretching

We will first consider the simple case of the TWLC ($G = 0$), following closely the approach of Ref. [1]. At high forces, the molecule is strongly oriented along the force direction, which is chosen to be parallel to $\widehat{\mathbf{z}}$. It proves convenient to decompose the tangent vector as

$$\widehat{\mathbf{e}}_3 = t_z\widehat{\mathbf{z}} + \mathbf{u}, \tag{A1}$$

where the vector $\mathbf{u}$ is orthogonal to $\widehat{\mathbf{z}}$, i.e. $\mathbf{u} = u_x\widehat{\mathbf{x}} + u_y\widehat{\mathbf{y}}$. Using the identity $|\widehat{\mathbf{e}}_3| = 1 = \sqrt{t_z^2 + \mathbf{u}^2}$ and expanding to lowest order in $\mathbf{u}$ we get

$$\widehat{\mathbf{e}}_3 = \left[1 - \frac{\mathbf{u}^2}{2} + \mathcal{O}(\mathbf{u}^4)\right]\widehat{\mathbf{z}} + \mathbf{u}, \tag{A2}$$

while its derivative is found to be

$$\frac{\mathrm{d}\widehat{\mathbf{e}}_3}{\mathrm{d}s} = \frac{\mathrm{d}\mathbf{u}}{\mathrm{d}s} + \mathcal{O}(\mathbf{u}^2)\widehat{\mathbf{z}}. \tag{A3}$$

Combining this with Eq. (1), we find

$$\Omega_1^2 + \Omega_2^2 = \left(\frac{\mathrm{d}\widehat{\mathbf{e}}_3}{\mathrm{d}s}\right)^2 = \left(\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}s}\right)^2 + \mathcal{O}(\mathbf{u}^3). \tag{A4}$$

Introducing the Fourier transform $\mathbf{u}_q = \int_0^L \mathrm{d}s\, e^{-iqs}\,\mathbf{u}(s)$, and neglecting higher-order terms, we write the bending and stretching contribution to the energy as follows

$$\frac{A}{2}\int_0^L \mathrm{d}s \left(\frac{\mathrm{d}\widehat{\mathbf{e}}_3}{\mathrm{d}s}\right)^2 - \beta\mathbf{f}\cdot\mathbf{R} = \frac{1}{2L}\sum_q \left(Aq^2 + \beta f\right)|\mathbf{u}_q|^2 - \beta f L, \tag{A5}$$

where we expressed the force as $\mathbf{f} = f\widehat{\mathbf{z}}$, while the end-to-end vector $\mathbf{R} = \int_0^L \mathrm{d}s\,\widehat{\mathbf{e}}_3$ was approximated based on Eq. (A2).

The torque in Eq. (5) is coupled to the linking number, which is the sum of twist and writhe [Eqs. (7) and (8), respectively]. In the high-force limit, Eq. (8) becomes

$$\begin{aligned} 2\pi\mathrm{Wr} &= \frac{1}{2}\int_0^L \left(\mathbf{u}\times\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}s}\right)\cdot\widehat{\mathbf{z}}\,\mathrm{d}s + \mathcal{O}(\mathbf{u}^4) \\ &= \frac{1}{2L}\sum_q \mathbf{u}_q^{\mathrm{T}}\begin{pmatrix} 0 & -iq \\ iq & 0\end{pmatrix}\mathbf{u}_q^* + \mathcal{O}(\mathbf{u}^4), \end{aligned} \tag{A6}$$

where we have rewritten the cross product as a matrix multiplication. Thus, the writhe couples the x and y components of the two-dimensional vector $\mathbf{u}_q$.

Adding up all terms, and with the help of simple algebraic manipulations, we obtain the following energy for the TWLC to lowest order in $\mathbf{u}$

$$\begin{aligned} &\beta E_0 - \beta\mathbf{f}\cdot\mathbf{R} - \beta\tau 2\pi\Delta\mathrm{Lk} \\ &= \frac{1}{2L}\sum_q \mathbf{u}_q^{\mathrm{T}}\mathbf{M}_q\mathbf{u}_q^* + \frac{C}{2}\int_0^L \mathrm{d}s\,\omega_3^2 - \frac{\beta^2\tau^2 L}{2C} - \beta f L, \end{aligned} \tag{A7}$$

where we introduced the matrix

$$\mathbf{M}_q = \begin{pmatrix} Aq^2 + \beta f & -iq\beta\tau \\ iq\beta\tau & Aq^2 + \beta f\end{pmatrix}. \tag{A8}$$

We have also introduced the shifted twist density

$$\omega_3(s) \equiv \Omega_3(s) - \frac{\beta\tau}{C}, \tag{A9}$$

which allowed us to eliminate linear terms in Ω_3. Thus, in the high-force limit, the TWLC under applied torque reduces, to lowest order, to a Gaussian model, where bending ($\mathbf{u}_q$) and twist (ω_3) are independent variables. The torque τ couples to the bending degrees of freedom through the off-diagonal terms of the matrix $\mathbf{M}_q$. The eigenvalues of $\mathbf{M}_q$ are easily found to be

$$\lambda_q^{\pm} = Aq^2 + \beta f \pm q\beta\tau, \tag{A10}$$

and the corresponding eigenvectors are $(\widehat{\mathbf{x}} \pm i\widehat{\mathbf{y}})/\sqrt{2}$. Writing Eq. (A7) on this basis allows us to calculate the partition function, from which the free energy is found to be

$$F_0 = -fL - \frac{\beta\tau^2 L}{2C} + \frac{k_{\mathrm{B}}T}{2}\sum_q \log\left(\lambda_q^+\lambda_q^-\right), \tag{A11}$$

where we have neglected additive constants. Expanding to quadratic order in τ

$$\log\left(\lambda_q^+\lambda_q^-\right) \approx \log\left(Aq^2 + \beta f\right)^2 - \left(\frac{q\beta\tau}{Aq^2 + \beta f}\right)^2, \tag{A12}$$

bioRxiv preprint doi: https://doi.org/10.1101/422683; this version posted September 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

and replacing the sum over momenta with an integral $\sum_q \to (L/2\pi)\int \mathrm{d}q$, we obtain

$$F_0(f,\tau) = F_0(f,0) - \frac{\beta\tau^2 L}{2C} - \frac{\beta\tau^2 L}{8A}\sqrt{\frac{k_\mathrm{B}T}{fA}} = F_0(f,0) - \frac{\beta\tau^2 L}{2C_\mathrm{eff}}. \tag{A13}$$

Combining the two last terms in the right-hand side, one obtains the Moroz and Nelson relation [Eq. (10)]. $F_0(f,\tau=0)$ is the zero-torque free energy, and is obtained by integrating the first term at the right-hand side of Eq. (A12). Although the integral is divergent, it can be regularized by introducing a momentum cutoff $\Lambda \approx 2\pi/a$, where $a = 0.34$ nm is the separation between neighboring base pairs. As it turns out, however, this cutoff does not affect any force-dependent terms, and one has

$$F_0(f,0) = F_0(0,0) - fL + \sqrt{\frac{fk_\mathrm{B}T}{A}}L, \tag{A14}$$

where $F_0(0,0)$ is cutoff-dependent. Interestingly, from Eq. (5) one finds to lowest order in τ

$$2\pi\Delta\mathrm{Lk} = -\frac{\partial F_0(f,\tau)}{\partial\tau} = \frac{\beta\tau L}{C_\mathrm{eff}}, \tag{A15}$$

which quantifies the induced over- or undertwisting upon the application of a torque. Using this expression, one obtains the force-extension relation at fixed linking number [4, 21]

$$\frac{z}{L} = -\frac{1}{L}\frac{\partial F_0}{\partial f} = 1 - \sqrt{\frac{k_\mathrm{B}T}{4fA}} - \frac{C^2}{2}\left(\frac{k_\mathrm{B}T}{4Af}\right)^{3/2}\left(\frac{2\pi\Delta\mathrm{Lk}}{L}\right)^2, \tag{A16}$$

which shows a characteristic parabolic profile for the extension of an over- or undertwisted, stretched molecule.

Appendix B: Ω_2 at strong stretching

The thermal average in Eq. (12) contains Ω_2, which needs to be expressed in terms of $\mathbf{u}_q$ and ω_3, the degrees of freedom of the system [Eq. (A7)]. For this purpose, we use the relation

$$\Omega_2 = \widehat{\mathbf{e}}_1 \cdot \frac{\mathrm{d}\widehat{\mathbf{e}}_3}{\mathrm{d}s}, \tag{B1}$$

which can be easily obtained from Eq. (1). In the high-force limit, where the tangent $\widehat{\mathbf{e}}_3$ points predominantly along the force direction, $\widehat{\mathbf{z}}$, one has

$$\widehat{\mathbf{e}}_1 = \left[\cos\psi + \mathcal{O}(\mathbf{u}^2)\right]\widehat{\mathbf{x}} + \left[\sin\psi + \mathcal{O}(\mathbf{u}^2)\right]\widehat{\mathbf{y}} + \mathcal{O}(\mathbf{u})\,\widehat{\mathbf{z}}. \tag{B2}$$

Here we have introduced the twist angle

$$\psi(s) = \int_0^s \Omega_3(t)\,\mathrm{d}t + \omega_0 s = \int_0^s \omega_3(t)\,\mathrm{d}t + \left(\omega_0 + \frac{\beta\tau}{C}\right)s, \tag{B3}$$

and used Eq. (A9) to express it in terms of the variable ω_3. Equation (B2) can be obtained by considering an arbitrary rotation that maps a fixed lab frame triad, e.g. $\{\widehat{\mathbf{x}},\widehat{\mathbf{y}},\widehat{\mathbf{z}}\}$, onto the material frame triad $\{\widehat{\mathbf{e}}_1,\widehat{\mathbf{e}}_2,\widehat{\mathbf{e}}_3\}$ at position s, requiring that $\widehat{\mathbf{e}}_3$ remains predominantly oriented along the force direction, as in Eq. (A1). Combining Eqs. (B1) and (B2), it follows that

$$\Omega_2 = \widehat{\mathbf{\Psi}}(s)\cdot\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}s} + \mathcal{O}(\mathbf{u}^3), \tag{B4}$$

where we have defined the unit vector

$$\widehat{\mathbf{\Psi}}(s) \equiv \cos\psi(s)\,\widehat{\mathbf{x}} + \sin\psi(s)\,\widehat{\mathbf{y}}. \tag{B5}$$

Therefore, in the high-force limit, Ω_2 can be written as a scalar product between a unit vector $\widehat{\mathbf{\Psi}}(s)$, depending exclusively on twist variables, and a vector $\mathrm{d}\mathbf{u}/\mathrm{d}s$, involving only the bending degrees of freedom. Finally, from Eq. (B4) it follows that

$$\widetilde{\Omega}_{2,q} = \frac{1}{L}\sum_{q'}(-iq')\mathbf{\Psi}_{q-q'}\cdot\mathbf{u}_{q'}. \tag{B6}$$

The remainder of the calculation, presented below, will be based upon Eqs. (B4) and (B6).

Appendix C: Details of the perturbative calculation

To calculate the average appearing in Eq. (12), it first needs to be rewritten as a function of the integration variable ω_3 [see Eq. (A9)]. This can be performed as follows

$$\begin{aligned}\left\langle\left(\int_0^L \mathrm{d}s\,\Omega_2\Omega_3\right)^2\right\rangle_0 &= \\ &= \frac{\beta^2\tau^2}{C^2}\left\langle\left(\int_0^L \mathrm{d}s\,\Omega_2\right)^2\right\rangle + \left\langle\left(\int_0^L \mathrm{d}s\,\Omega_2\omega_3\right)^2\right\rangle \\ &= \frac{\beta^2\tau^2}{C^2}\left\langle\widetilde{\Omega}_{2,0}^2\right\rangle + \frac{1}{L^2}\sum_{q,k}\left\langle\widetilde{\Omega}_{2,q}\widetilde{\Omega}_{2,k}\widetilde{\omega}_{3,-q}\widetilde{\omega}_{3,-k}\right\rangle,\end{aligned} \tag{C1}$$

where $\widetilde{\Omega}_{2,q}$ and $\widetilde{\omega}_{3,q}$ denote the Fourier components of Ω_2 and ω_3, respectively. Note that we have neglected a linear term in ω_3, which vanishes due to the symmetry $\omega_3 \leftrightarrow -\omega_3$. Moreover, in order to simplify the notation, we have dropped the subscript from all averages $\langle.\rangle_0$, which will be always calculated within the TWLC model, i.e. for $G = 0$.

Before proceeding to the calculation of Eq. (C1), it will prove useful to first present some properties. In particular, we are going to use the following expressions, obtained from the correlation functions in the TWLC model [Eq. (A7)]

$$\langle\mathbf{u}_q\cdot\mathbf{u}_k\rangle = \frac{2L\left(Aq^2+\beta f\right)}{\left(Aq^2+\beta f\right)^2 - \left(q\beta\tau\right)^2}\,\delta_{q,-k} \tag{C2}$$

bioRxiv preprint doi: https://doi.org/10.1101/422683; this version posted September 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

and

$$\langle \mathbf{u}_q \otimes \mathbf{u}_k \rangle = \frac{2iLq\beta\tau}{(Aq^2+\beta f)^2 - (q\beta\tau)^2}\,\delta_{q,-k}. \tag{C3}$$

For convenience, we have introduced the shorthand notation

$$\mathbf{a} \otimes \mathbf{b} \equiv \widehat{\mathbf{z}} \cdot (\mathbf{a} \times \mathbf{b}) = a_x b_y - a_y b_x, \tag{C4}$$

which is antisymmetric with respect to the interchange of $\mathbf{a}$ and $\mathbf{b}$. From Eq. (C3) it follows that $\langle \mathbf{u}_q \otimes \mathbf{u}_k \rangle = 0$, when $\tau = 0$ (in this case the matrix $\mathbf{M}_q$ is diagonal, hence the cross-correlations $\langle u_q^x u_{-q}^y \rangle = 0$). Moreover, for $\tau = 0$ and $q = -k$, Eq. (C2) reduces to:

$$\langle |\mathbf{u}_q|^2 \rangle_{\tau=0} = \frac{2L}{Aq^2+\beta f}, \tag{C5}$$

which can be easily obtained from equipartition [12]. We are also going to use the following symmetries

$$\langle u_q^x u_k^x \rangle = \langle u_q^y u_k^y \rangle, \tag{C6}$$
$$\langle u_q^x u_k^y \rangle = -\langle u_q^y u_k^x \rangle, \tag{C7}$$

which allow us to rearrange scalar products as follows

$$\begin{aligned}&\langle \mathbf{\Psi}_{-q} \cdot \mathbf{u}_q \mathbf{\Psi}_{-k} \cdot \mathbf{u}_k \rangle = \langle u_q^x u_k^x \rangle \left[\langle \Psi_{-q}^x \Psi_{-k}^x \rangle + \right.\\ &\left. \langle \Psi_{-q}^y \Psi_{-k}^y \rangle \right] + \langle u_q^x u_k^y \rangle \left[\langle \Psi_{-q}^x \Psi_{-k}^y \rangle - \langle \Psi_{-q}^y \Psi_{-k}^x \rangle \right]\\ &= \frac{1}{2}\left[\langle \mathbf{u}_q \cdot \mathbf{u}_k \rangle \langle \mathbf{\Psi}_{-q} \cdot \mathbf{\Psi}_{-k} \rangle + \langle \mathbf{u}_q \otimes \mathbf{u}_k \rangle \langle \mathbf{\Psi}_{-q} \otimes \mathbf{\Psi}_{-k} \rangle \right],\end{aligned} \tag{C8}$$

where we have used the fact that the bending ($\mathbf{u}$) and twisting ($\mathbf{\Psi}$) degrees of freedom are independent, within the TWLC. We are now ready to proceed to the calculation of Eq. (C1). We will need to evaluate two distinct terms, which will be treated separately.

1. First term in Eq. (C1)

The first term in Eq. (C1) already contains a factor of order $\mathcal{O}(\tau^2)$, which means that up to quadratic order in τ it is sufficient to evaluate the corresponding average for $\tau = 0$

$$\begin{aligned}\left\langle \widetilde{\Omega}_{2,0}^2 \right\rangle_{\tau=0} &= \frac{1}{L^2}\sum_{qk}(-qk)\langle \mathbf{\Psi}_{-q} \cdot \mathbf{u}_q \mathbf{\Psi}_{-k} \cdot \mathbf{u}_k \rangle_{\tau=0}\\ &= \frac{1}{2L^2}\sum_q q^2 \langle |\mathbf{u}_q|^2 \rangle_{\tau=0} \langle |\mathbf{\Psi}_q|^2 \rangle_{\tau=0},\end{aligned} \tag{C9}$$

where we have used Eqs. (B6) and (C8), together with the property $\langle \mathbf{u}_q \otimes \mathbf{u}_k \rangle_{\tau=0} = 0$ [see Eq. (C3)]. Next, we need to calculate the following quantity

$$\langle |\mathbf{\Psi}_q|^2 \rangle_{\tau=0} = \int_0^L \mathrm{d}s\mathrm{d}s' e^{iq(s-s')} \langle \widehat{\mathbf{\Psi}}(s) \cdot \widehat{\mathbf{\Psi}}(s') \rangle_{\tau=0}. \tag{C10}$$

From Eqs. (B3) and (B5) one finds

$$\begin{aligned}\langle \widehat{\mathbf{\Psi}}(s) \cdot \widehat{\mathbf{\Psi}}(s') \rangle_{\tau=0} &= \langle \cos[\psi(s) - \psi(s')] \rangle\\ &= \frac{e^{i\omega_0(s'-s)}}{2}\left\langle \exp\left(i\int_s^{s'} \omega_3(t)\,\mathrm{d}t \right) \right\rangle + \text{c.c.},\end{aligned} \tag{C11}$$

where c.c. denotes the complex conjugate. To proceed, we perform a Fourier transform of the exponent

$$\int_s^{s'} \omega_3(t)\,\mathrm{d}t = \frac{1}{L}\sum_q h_q\,\widetilde{\omega}_{3,q}, \tag{C12}$$

where we have introduced the complex variable

$$h_q = \frac{e^{-iqs'} - e^{-iqs}}{-iq}. \tag{C13}$$

Performing Gaussian integration in $\widetilde{\omega}_{3,q}$, one finds

$$\begin{aligned}\left\langle \exp\left(\pm\frac{i}{L}\sum_q h_q\,\widetilde{\omega}_{3,q} \right) \right\rangle &= \exp\left(-\frac{1}{LC}\sum_{q>0} h_q h_{-q} \right)\\ &= \exp\left(-\frac{|s'-s|}{2C} \right).\end{aligned} \tag{C14}$$

As expected, the decay of the twist correlation function is governed by $2C$, i.e. the twist persistence length in the TWLC. Combining Eqs. (C11), (C12) and (C14), we obtain

$$\langle \widehat{\mathbf{\Psi}}(s) \cdot \widehat{\mathbf{\Psi}}(s') \rangle_{\tau=0} = \cos[\omega_0(s'-s)]\, e^{-|s'-s|/2C}. \tag{C15}$$

Inserting this in Eq. (C10) yields

$$\begin{aligned}\langle |\mathbf{\Psi}_q|^2 \rangle_{\tau=0} &= \frac{L}{2}\left(\frac{1}{i(q-\omega_0)+\frac{1}{2C}} + \frac{1}{i(q+\omega_0)+\frac{1}{2C}} \right.\\ &\quad \left. + \frac{1}{i(-q+\omega_0)+\frac{1}{2C}} + \frac{1}{i(-q-\omega_0)+\frac{1}{2C}} \right)\\ &= \frac{L}{2C}\left[\frac{1}{(q+\omega_0)^2+\frac{1}{4C^2}} + \frac{1}{(q-\omega_0)^2+\frac{1}{4C^2}} \right].\end{aligned} \tag{C16}$$

Finally, combining Eqs. (C9), (C5) and (C16) we find

$$\begin{aligned}\left\langle \widetilde{\Omega}_{2,0}^2 \right\rangle_{\tau=0} &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \frac{\mathrm{d}q\, q^2}{Aq^2+\beta f}\left\langle |\mathbf{\Psi}_q|^2 \right\rangle\\ &= \frac{L}{A}\left[1 - \sqrt{\frac{\beta f}{A}}\frac{\sqrt{\frac{\beta f}{A}}+\frac{1}{2C}}{\left(\sqrt{\frac{\beta f}{A}}+\frac{1}{2C}\right)^2+\omega_0^2} \right]\\ &\equiv \frac{L}{A}d(f),\end{aligned} \tag{C17}$$

where we introduced a force-dependent scale factor $d(f)$. Note that $0 \le d(f) \le 1$, with $d(f) \to 1$ at small forces and $d(f) \to 0$ at high forces. Commonly accepted estimates of the DNA elastic constants put them in the viccinity of $C = 100$ nm and $A = 50$ nm, while the

bioRxiv preprint doi: https://doi.org/10.1101/422683; this version posted September 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

applied forces in typical experiments are in the range 0.1 pN $\lesssim f \lesssim$ 10 pN. Recalling that room temperature corresponds to $k_{\rm B}T \approx 4$ pN·nm, it follows that $\beta f/A$ is at least one order of magnitude larger than $1/4C^2$. This allows for the following simplification

$$d(f) \approx \frac{1}{1+f/f_0}, \tag{C18}$$

neglecting higher-order terms in $C^{-1}\sqrt{A/\beta f}$. We have also introduced a characteristic force

$$f_0 = Ak_{\rm B}T\omega_0^2 \approx 600 \text{ pN}, \tag{C19}$$

whose value greatly exceeds those at which the double helix breaks. Thus, for the force range of interest, we may set $d(f) = 1$ in Eq. (C17). Summarizing, this first term of Eq. (C1) provides the following contribution to the free energy

$$\Delta F^{(1)} = -\frac{G^2}{2}\frac{\beta\tau^2}{C^2}\left\langle \widetilde{\Omega}_{2,0}^2 \right\rangle = -\frac{\beta\tau^2 L}{2C}\frac{G^2}{AC}d(f). \tag{C20}$$

As a final remark, we note that we could have obtained Eq. (C18) using the approximation

$$\langle |\mathbf{\Psi}_q|^2\rangle \approx \pi L[\delta(q+\omega_0)+\delta(q-\omega_0)]. \tag{C21}$$

Formally, this corresponds to taking the limit $C \to \infty$ in Eq. (C16), i.e. approximating the Lorentzian distributions $1/[2C(q\pm\omega_0)^2+1/2C]$ with delta functions. This is a valid approximation as long as $\omega_0 \gg 1/2C$, making the Lorentzians sharply-peaked at large momenta $q = \pm\omega_0$, where the integrand in Eq. (C17) varies slowly.

2. Second term in Eq. (C1)

Using the same decomposition as in Eq. (C8), the second term in Eq. (C1) can be written as

$$\begin{aligned}
&\sum_{q,k}\left\langle \widetilde{\Omega}_{2,q}\,\widetilde{\omega}_{3,-q}\,\widetilde{\Omega}_{2,k}\,\widetilde{\omega}_{3,-k}\right\rangle = \frac{1}{L^2}\sum_{q,k,q',k'}(-q'k')\langle \mathbf{\Psi}_{q-q'}\cdot\mathbf{u}_{q'}\,\mathbf{\Psi}_{k-k'}\cdot\mathbf{u}_{k'}\,\widetilde{\omega}_{3,-q}\widetilde{\omega}_{3,-k}\rangle \\
&= \frac{1}{2L^2}\sum_{q,k,q',k'}(-q'k')[\langle \mathbf{u}_{q'}\cdot\mathbf{u}_{k'}\rangle\langle \mathbf{\Psi}_{q-q'}\cdot\mathbf{\Psi}_{k-k'}\,\widetilde{\omega}_{3,-q}\widetilde{\omega}_{3,-k}\rangle + \langle \mathbf{u}_{q'}\otimes\mathbf{u}_{k'}\rangle\langle \mathbf{\Psi}_{q-q'}\otimes\mathbf{\Psi}_{k-k'}\,\widetilde{\omega}_{3,-q}\widetilde{\omega}_{3,-k}\rangle] \\
&= \frac{1}{2L^2}\sum_{q'}q'^2[\langle |\mathbf{u}_{q'}|^2\rangle I_{\rm s}(q',\tau) + \langle \mathbf{u}_{q'}\otimes\mathbf{u}_{-q'}\rangle I_{\rm a}(q',\tau)],
\end{aligned} \tag{C22}$$

where we used the fact that the $\mathbf{u}$ correlators are diagonal in momentum space, hence $k' = -q'$ [see Eqs. (C2) and (C3)]. We have also introduced the symmetric

$$I_{\rm s}(q',\tau) = \sum_{q,k}\langle \mathbf{\Psi}_{q-q'}\cdot\mathbf{\Psi}_{k+q'}\,\widetilde{\omega}_{3,-q}\,\widetilde{\omega}_{3,-k}\rangle, \tag{C23}$$

and antisymmetric products

$$I_{\rm a}(q',\tau) = \sum_{q,k}\langle \mathbf{\Psi}_{q-q'}\otimes\mathbf{\Psi}_{k+q'}\,\widetilde{\omega}_{3,-q}\,\widetilde{\omega}_{3,-k}\rangle. \tag{C24}$$

In what follows, we are going to compute the contribution of $I_{\rm s}$ and $I_{\rm a}$ to the free energy separately.

a. *Symmetric products*

For the evaluation of Eq. (C23), we will first focus on the average inside the summation, which may be written in the following way

$$\begin{aligned}
&\langle \mathbf{\Psi}_{q-q'}\cdot\mathbf{\Psi}_{k+q'}\widetilde{\omega}_{3,-q}\widetilde{\omega}_{3,-k}\rangle = \\
&\int_0^L {\rm d}s{\rm d}s'\, e^{i(q-q')s+i(k+q')s'}\langle \widehat{\mathbf{\Psi}}(s)\cdot\widehat{\mathbf{\Psi}}(s')\,\widetilde{\omega}_{3,-q}\widetilde{\omega}_{3,-k}\rangle.
\end{aligned} \tag{C25}$$

We may now use Eqs. (C11)-(C14) so as to obtain

$$\begin{aligned}
&\langle \widehat{\mathbf{\Psi}}(s)\cdot\widehat{\mathbf{\Psi}}(s')\,\widetilde{\omega}_{3,-q}\widetilde{\omega}_{3,-k}\rangle \\
&= -\frac{L^2}{2}\left[e^{i\overline{\omega}_0(s'-s)}\frac{\partial^2}{\partial h_{-q}\partial h_{-k}}\left\langle \exp\left(\frac{i}{L}\sum_p h_p\,\widetilde{\omega}_{3,p}\right)\right\rangle\right. \\
&\quad \left. + e^{-i\overline{\omega}_0(s'-s)}\frac{\partial^2}{\partial h_{-q}\partial h_{-k}}\left\langle \exp\left(-\frac{i}{L}\sum_p h_p\,\widetilde{\omega}_{3,p}\right)\right\rangle\right] \\
&= -L^2\cos\left[\overline{\omega}_0(s'-s)\right]\frac{\partial^2 e^{-\frac{1}{LC}\sum_{p>0}h_p h_{-p}}}{\partial h_{-q}\partial h_{-k}} \\
&= \frac{L}{C}\cos\left[\overline{\omega}_0(s'-s)\right]\left(\delta_{q,-k}-\frac{1}{LC}h_q h_k\right)e^{-|s'-s|/2C},
\end{aligned} \tag{C26}$$

bioRxiv preprint doi: https://doi.org/10.1101/422683; this version posted September 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

where we have introduced the shifted intrinsic twist

$$\overline{\omega}_0 \equiv \omega_0 + \frac{\beta\tau}{C}. \tag{C27}$$

Differently from the calculation of $\langle \Omega_{2,0}^2 \rangle$ in Eq. (C9), we can no longer ignore the torque dependence of ψ [see Eq. (B3)]. Plugging Eq. (C26) back in Eq. (C25), integrating in s and s' and summing over q and k, we obtain

$$\begin{aligned} I_{\rm s}(q',\tau) &= \frac{L^3}{C} - \frac{L^3}{4C^2}\int_{-\infty}^{+\infty} {\rm d}r \cos(q'r)\cos(\overline{\omega}_0 r)\, e^{-|r|/2C} \\ &= \frac{L^3}{C} - \frac{L^3}{4C^3}\left[\frac{1}{(q'+\overline{\omega}_0)^2+\frac{1}{4C^2}} + \frac{1}{(q'-\overline{\omega}_0)^2+\frac{1}{4C^2}}\right] \\ &\approx \frac{L^3}{C} - \frac{\pi L^3}{2C^2}\left[\delta\left(q+\overline{\omega}_0\right) + \delta\left(q-\overline{\omega}_0\right)\right]. \end{aligned} \tag{C28}$$

Throughout the calculation we introduced the variable $r \equiv s'-s$ in the double integral. Similar to Eq. (C21), we also approximated the two Lorentzians with delta functions. Note that $I_{\rm s}$ depends on the torque τ through $\overline{\omega}_0$, as indicated by Eq. (C27). Combining Eqs. (C2) and (C28), one finds

$$\begin{aligned} &\frac{1}{2L^2}\sum_q q^2 \left\langle \mathbf{u}_q \cdot \mathbf{u}_{-q}\right\rangle I_{\rm s}(q,\tau) \\ &= \frac{1}{L}\sum_q \frac{Aq^2+\beta f}{\left(Aq^2+\beta f\right)^2 - \left(q\beta\tau\right)^2} q^2 I_{\rm s}(q,\tau) \\ &= \frac{1}{L}\sum_q \frac{q^2 I_{\rm s}(q,\tau)}{Aq^2+\beta f} + \frac{\beta^2\tau^2}{L}\sum_q \frac{q^4 I_{\rm s}(q,0)}{(Aq^2+\beta f)^3} + \mathcal{O}(\tau^4). \end{aligned} \tag{C29}$$

We are interested in terms proportional to τ^2. There are two such contributions, the first one being

$$\begin{aligned} &\frac{1}{L}\sum_q \frac{q^2 I_{\rm s}(q,\tau)}{Aq^2+\beta f} = \frac{-L^3}{AC}\left(\sqrt{\frac{\beta f}{4A}} + \frac{1}{2C}\frac{\overline{\omega}_0^2}{A\overline{\omega}_0^2+\beta f} + \dots\right) \\ &= -\frac{L^3}{AC}\sqrt{\frac{\beta f}{4A}} - \frac{L^3 d(f)}{2AC^2} \times \\ &\quad \times \left\{1 + \frac{\beta\tau}{C\omega_0}\widetilde{d}(f) + \frac{\beta^2\tau^2}{C^2\omega_0^2}\widetilde{d}(f)[4d(f)-1]\right\} + \dots \\ &= -\frac{L^3}{AC}\sqrt{\frac{\beta f}{4A}} + \dots, \end{aligned} \tag{C30}$$

where we defined $\widetilde{d}(f) \equiv 1 - d(f)$, with $d(f)$ the scale factor given in Eq. (C18), and where the dots indicate omitted terms, which do not significantly contribute to the result. These terms are either independent of the torque and force, or are of higher order than τ^2. We note that $\widetilde{d}(f) \ll 1$, i.e. it is negligibly small for the experimentally-accessible forces $f \ll f_0 \approx 600$ pN. The only surviving term in Eq. (C30) is independent of τ and proportional to $\sqrt{f}$, hence contributing to the force-extension response. The remaining term to evaluate in Eq. (C29) is

$$\begin{aligned} &\frac{\beta^2\tau^2}{L}\sum_q \frac{q^4 I_{\rm s}(q,0)}{(Aq^2+\beta f)^3} = \frac{\beta^2\tau^2}{4\pi}\frac{\partial^2}{\partial A^2}\int \frac{{\rm d}q}{Aq^2+\beta f} I_{\rm s}(q,0) \\ &= \frac{\beta^2\tau^2 L^3}{C}\left[\frac{3}{16A^2}\sqrt{\frac{k_{\rm B}T}{fA}} - \frac{2d^3(f)}{CA^3\omega_0^2}\right] \\ &\approx \frac{3\beta^2\tau^2 L^3}{16A^2C}\sqrt{\frac{k_{\rm B}T}{fA}}. \end{aligned} \tag{C31}$$

Note that terms containing ω_0^{-2} are always multiplied by A^{-2} or C^{-2}, hence forming dimensionless constants. Typical values for the case of DNA are $(A\omega_0)^{-2} \approx 10^{-4}$ and $(C\omega_0)^{-2} \approx 3\times 10^{-5}$, which provide negligible contributions to the free energy, compared to other terms of the same order in τ. Therefore, the term proportional to d^3 in Eq. (C31) can be safely neglected. Combining Eqs. (C29)-(C31), we find that the relevant contribution of the symmetric term in Eq. (C22) to the free energy is

$$\Delta F^{(2)} = \frac{G^2 L}{4AC}\sqrt{\frac{f k_{\rm B}T}{A}} - \frac{\beta\tau^2 L}{2C}\frac{3G^2}{16A^2}\sqrt{\frac{k_{\rm B}T}{fA}}. \tag{C32}$$

b. *Antisymmetric products*

The final part of the derivation is devoted to the calculation of the antisymmetric product in Eq. (C22). We start by expanding Eq. (C3) as follows

$$\left\langle \mathbf{u}_q \otimes \mathbf{u}_{-q}\right\rangle = \frac{-2iLq\beta\tau}{\left(Aq^2+\beta f\right)^2} + \mathcal{O}(\tau^3). \tag{C33}$$

The calculation of the twist correlator is performed in a similar fashion as above, which yields

$$\begin{aligned} &\langle \widehat{\mathbf{\Psi}}(s) \otimes \widehat{\mathbf{\Psi}}(s')\, \widetilde{\omega}_{3,-q}\widetilde{\omega}_{3,-k}\rangle = \\ &\frac{L}{C}\sin\left[\overline{\omega}_0(s'-s)\right]\left(\delta_{q,-k} - \frac{1}{LC}h_q h_k\right) e^{-|s'-s|/2C}. \end{aligned} \tag{C34}$$

We may take the Fourier transform of this expression, and plug it back into Eq. (C24), so as to obtain

$$\begin{aligned} I_{\rm a}(q',\tau) &= -\frac{iL^3}{4C^3}\left[\frac{1}{(q'+\overline{\omega}_0)^2+\frac{1}{4C^2}} - \frac{1}{(q'-\overline{\omega}_0)^2+\frac{1}{4C^2}}\right] \\ &\approx -\frac{i\pi L^3}{2C^2}\left[\delta\left(q+\overline{\omega}_0\right) - \delta\left(q-\overline{\omega}_0\right)\right]. \end{aligned} \tag{C35}$$

Finally, plugging Eqs. (C33) and (C35) into the second term of Eq. (C22), transforming the sum into an integral

bioRxiv preprint doi: https://doi.org/10.1101/422683; this version posted September 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

and performing the remaining integration, we find

$$\begin{aligned}
&\frac{1}{2L^2}\sum_q q^2 \left\langle \mathbf{u}_q \otimes \mathbf{u}_{-q}\right\rangle I_{\mathrm{a}}(q,\tau) \\
&\approx -\frac{i}{L}\sum_q \frac{q\beta\tau}{\left(Aq^2+\beta f\right)^2} q^2 I_{\mathrm{a}}(q,\tau) \\
&= \frac{i\beta\tau}{2\pi}\frac{\partial}{\partial A}\int \frac{\mathrm{d}q q I_{\mathrm{a}}(q,\tau)}{Aq^2+\beta f} = \frac{\beta\tau L^3}{2C^2}\frac{\partial}{\partial A}\frac{\overline{\omega}_0}{A\overline{\omega}_0^2+\beta f} \\
&= -\frac{\beta\tau L^3}{2C^2}\frac{\overline{\omega}_0^3}{\left(A\overline{\omega}_0^2+\beta f\right)^2} \approx -\frac{\beta\tau d^2(f)L^3}{2A^2C^2\omega_0},
\end{aligned} \tag{C36}$$

where we have omitted terms, which are either higher order in τ, or negligibly small compared to other terms of the same order [recall $(C\omega_0)^{-2} \approx 3\times 10^{-5}$]. Summarizing, the contribution of the antisymmetric product to the free energy is

$$\Delta F^{(3)} = \frac{G^2 d^2(f)\tau L}{4A^2C^2\omega_0}. \tag{C37}$$

3. Collecting the results

Throughout the derivation we found three distinct contributions to the free energy, coming from Eqs. (C20), (C32) and (C37). Adding these to Eq. (A13), i.e. the free energy of the TWLC, we find

$$\frac{1}{L}F(f,\tau) \approx -f + \sqrt{\frac{fk_{\mathrm{B}}T}{A_{\mathrm{eff}}}} + \Gamma\tau - \frac{\beta\tau^2}{2C_{\mathrm{eff}}}, \tag{C38}$$

where we have omitted both terms independent of f and τ and higher-order corrections in τ and G. We have also introduced the effective bending stiffness

$$\frac{1}{A_{\mathrm{eff}}} = \frac{1}{A}\sqrt{1+\frac{G^2}{4AC}} \approx \frac{1}{A}\left(1+\frac{G^2}{2AC}\right), \tag{C39}$$

together with the proportionality constant

$$\Gamma = \frac{G^2 d^2(f)}{4A^2C^2\omega_0}. \tag{C40}$$

Finally, we reach the following expression for the effective torsional stiffness

$$\frac{1}{C_{\mathrm{eff}}} = \frac{1}{C}\left[1+\frac{G^2}{AC}d(f)\right] + \frac{1}{4A}\left(1+\frac{3G^2}{4AC}\right)\sqrt{\frac{k_{\mathrm{B}}T}{fA}}, \tag{C41}$$

corresponding to Eq. (16) of the main text, and the central result of this work.

Appendix D: Intrinsic bending

The analysis above is based on description of the ground-state configuration of DNA relative to a straight molecular axis [$\mathbf{\Omega} = \mathbf{0}$ in Eq. (1)], i.e., for a molecular axis which is straight in the ground state. However, one can also choose coordinates where the ground state of the double helix is a helix while still respecting the symmetry of the elastic model. In fact, this is a rather natural outcome for most choices of DNA deformation which are based on molecular modeling, where coordinates are usually chosen relative to the orientations of the base pairs (e.g., using the vector connecting the junctions of the bases to the sugar-phosphate backbone as a reference), due to the groove asymmetry of DNA. Most relevant here, our previous determination of the elastic constants of oxDNA2 [15] analyzed deformations relative to a helical coordinate system. We now show how to transform the elastic constants in such a helical coordinate system to the straight-line coordinates relevant to our calculations.

Intrinsic bending consistent with groove asymmetry, usually reported in the DNA literature as a nonzero value of the average roll [37], can be described using the following modification of Eq. (1)

$$\frac{\mathrm{d}\widehat{\mathbf{e}}_i}{\mathrm{d}s} = \left(\mathbf{\Omega} + l_2\widehat{\mathbf{e}}_2 + l_3\widehat{\mathbf{e}}_3\right)\times\widehat{\mathbf{e}}_i, \tag{D1}$$

where l_2 and l_3 correspond to the intrinsic bending and twisting densities, respectively, with $l_2 \ll l_3$. A nonzero l_1 is incompatible with the symmetry of the double helix.

Solving Eq. (D1) for $\mathbf{\Omega} = \mathbf{0}$, one finds that the ground-state configuration is a helix, with a linking number equal to $\mathrm{Lk}_0 = \omega_0 L/2\pi$, where

$$\omega_0 = \sqrt{l_2^2 + l_3^2}. \tag{D2}$$

Furthermore, from the solution of Eq. (D1), it follows that the rotation matrix transforming the helical ground state of Eq. (D1) to the straight one of Eq. (1) is

$$\mathbf{R} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & l_3/\omega_0 & l_2/\omega_0 \\ 0 & -l_2/\omega_0 & l_3/\omega_0 \end{pmatrix}, \tag{D3}$$

expressed on the body frame $\{\widehat{\mathbf{e}}_1, \widehat{\mathbf{e}}_2, \widehat{\mathbf{e}}_3\}$ of the former.

The total elastic free energy should not depend on the coordinate system used to describe it, so the energy in the helical coordinates ($l_2, l_3 \neq 0$) should equal that found in non-helical coordinates ($l_2 = 0$ and $l_3 = \omega_0$) The deformations in the "straight" model are given by $\mathbf{\Omega}' = \mathbf{R}^{\mathrm{T}}\mathbf{\Omega}$, where $\mathbf{\Omega}$ are the deformation parameters of the helical model. From the condition that the integrand of Eq. (2) has to remain invariant under this transformation, one obtains the following relations mapping the elastic con-

bioRxiv preprint doi: https://doi.org/10.1101/422683; this version posted September 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

stants from the helical coordinates to the straight ones:

$$A_{1,\mathrm{s}} = A_1, \tag{D4}$$

$$A_{2,\mathrm{s}} = A_2 - \frac{2xG - x^2(C - A_2)}{1 + x^2}, \tag{D5}$$

$$C_{\mathrm{s}} = C + \frac{2xG - x^2(C - A_2)}{1 + x^2}, \tag{D6}$$

$$G_{\mathrm{s}} = G - \frac{x(C - A_2) + 2x^2 G}{1 + x^2}, \tag{D7}$$

where $x \equiv l_2/l_3 \ll 1$, and where the subscript s indicates the "straight" frame result. The transformation (D3) mixes A_2, G and C, changing their values, but conserves the symmetry of the elastic constant matrix. These formulae allow one to measure elastic constants using arbitrarily chosen helical reference coordinates, and then convert them to elastic constants suitable for using strains defined relative to a straight-line ground state.

For unconstrained (zero force and torque) oxDNA2 simulations, we measured reference helix parameters $l_2 = 0.1349$ nm^{-1} and $l_3 = 1.774$ nm^{-1}, giving $\omega_0 = 1.779$ nm^{-1} and $x = 0.076$. Elastic constants reported in Ref. [15] ($A_1 = 85$ nm, $A_2 = 39$ nm, $C = 105$ nm, and $G = 30$ nm) were measured in reference to helical coordinates; for use in our analytical theory we transform them to the straight coordinates using (D4)-(D7) to obtain the oxDNA2 values in Table 1.

[1] J. F. Marko, "Biophysics of protein-DNA interactions and chromosome organization," Physica A **418**, 126–153 (2015).

[2] C Bustamante, J Marko, E Siggia, and B Smith, "Entropic elasticity of λ-phage DNA," Science **265**, 1599–1599 (1994).

[3] TR Strick, J-F Allemand, D Bensimon, A Bensimon, and V Croquette, "The elasticity of a single supercoiled DNA molecule," Science **271**, 1835–1837 (1996).

[4] J. D. Moroz and P. Nelson, "Torsional directed walks, entropic elasticity, and DNA twist stiffness," Proc. Natl. Acad. Sci. USA **94**, 14418–14422 (1997).

[5] C. Bouchiat and M. Mézard, "Elasticity model of a supercoiled DNA molecule," Phys. Rev. Lett. **80**, 1556–1559 (1998).

[6] F. Mosconi, J. F. Allemand, D. Bensimon, and V. Croquette, "Measurement of the torque on a single stretched and twisted DNA using magnetic tweezers," Phys. Rev. Lett. **102**, 078301 (2009).

[7] T. Guérin, "Analytical expressions for the closure probability of a stiff wormlike chain for finite capture radius," Phys. Rev. E **96**, 022501 (2017).

[8] Céline Barde, Nicolas Destainville, and Manoel Manghi, "Energy required to pinch a DNA plectoneme," Phys. Rev. E **97**, 032412 (2018).

[9] K. D. Whitley, M. J. Comstock, and Y. R. Chemla, "Ultrashort nucleic acid duplexes exhibit long wormlike chain behavior with force-dependent edge effects," Phys. Rev. Lett. **120**, 068102 (2018).

[10] M. Liangruksa and S. Wongwises, "An elastic model of DNA under thermal induced stress," Math. Biosciences **300**, 47–54 (2018).

[11] T. Bleha and P. Cifra, "Correlation anisotropy and stiffness of DNA molecules confined in nanochannels," J. Chem. Phys. **149**, 054903 (2018).

[12] J. F. Marko and E. D. Siggia, "Bending and twisting elasticity of DNA," Macromolecules **27**, 981–988 (1994).

[13] F. Mohammad-Rafiee and R. Golestanian, "Elastic correlations in nucleosomal DNA structure," Phys. Rev. Lett. **94**, 238102 (2005).

[14] S. K. Nomidis, F. Kriegel, W. Vanderlinden, J. Lipfert, and E. Carlon, "Twist-Bend Coupling and the Torsional Response of Double-Stranded DNA," Phys. Rev. Lett. **118**, 217801 (2017).

[15] E. Skoruppa, M. Laleman, S. Nomidis, and E. Carlon, "DNA elasticity from coarse-grained simulations: The effect of groove asymmetry," J. Chem. Phys. **146**, 214902 (2017).

[16] E. Skoruppa, S. K. Nomidis, J. F. Marko, and E. Carlon, "Bend-Induced Twist Waves and the Structure of Nucleosomal DNA," Phys. Rev. Lett. **121**, 088101 (2018).

[17] J. Lipfert, J. Kerssemakers, T. Jager, and N. Dekker, "Magnetic torque tweezers: measuring torsional stiffness in DNA and RecA-DNA filaments," Nat. Methods **7**, 977–980 (2010).

[18] D. J. Kauert, T. Kurth, T. Liedl, and R. Seidel, "Direct mechanical measurements reveal the material properties of three-dimensional DNA origami," Nano Lett. **11**, 5558–5563 (2011).

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

[20] J. Lipfert, G. M. Skinner, J. M. Keegstra, T. Hensgens, T. Jager, D. Dulin, M. Köber, Z. Yu, S. P. Donkers, F.-C. Chou, R. Das, and N. H. Dekker, "Double-stranded RNA under force and torque: Similarities to and striking differences from double-stranded DNA," Proc. Natl. Acad. Sci. USA **111**, 15408–15413 (2014).

[21] J. D. Moroz and P. Nelson, "Entropic elasticity of twist-storing polymers," Macromolecules **31**, 6333–6347 (1998).

[22] J. Lipfert, M. Wiggin, J. W. J. Kerssemakers, F. Pedaci, and N. H. Dekker, "Freely orbiting magnetic tweezers to directly monitor changes in the twist of nucleic acids," Nat. Commun. **2**, 439 (2011).

[23] F. B. Fuller, "Decomposition of the linking number of a closed ribbon: A problem from molecular biology." Proc. Natl. Acad. Sci. USA **75**, 3557–3561 (1978).

[24] Filip Lankaš, Jiri Šponer, Jorg Langowski, and Thomas E Cheatham, "DNA basepair step deformability inferred from molecular dynamics simulations." Biophys J. **85**, 2872–2883 (2003).

[25] B. Eslami-Mossallam and M. R. Ejtehadi, "Asymmetric elastic rod model for DNA," Phys. Rev. E **80**, 011919 (2009).

bioRxiv preprint doi: https://doi.org/10.1101/422683; this version posted September 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

[26] D. E. Depew and J. C. Wang, "Conformational fluctuations of DNA helix," Proc. Natl. Acad. Sci. USA **72**, 4275–4279 (1975).

[27] M. Duguet, "The helical repeat of DNA at high temperature," Nucl. Acids Res. **21**, 463–468 (1993).

[28] F. Kriegel, C. Matek, T. Dršata, K. Kulenkampff, S. Tschirpke, M. Zacharias, F. Lankaš, and J. Lipfert, "The temperature dependence of the helical twist of DNA," Nucl. Acids Res. **46**, 7998–8009 (2018).

[29] T. E. Ouldridge, A. A. Louis, and J. P. K. Doye, "DNA nanotweezers studied with a coarse-grained model of DNA," Phys. Rev. Lett. **104**, 178101 (2010).

[30] F.-C. Chou, J. Lipfert, and R. Das, "Blind predictions of dna and rna tweezers experiments with force and torque," PLoS Comp. Biol. **10**, e1003756 (2014).

[31] P. Šulc, F. Romano, T. Ouldridge, L. Rovigatti, J. Doye, and A. Louis, "Sequence-dependent thermodynamics of a coarse-grained DNA model," J. Chem. Phys. **137**, 135101 (2012).

[32] B. Snodin, F. Randisi, M. Mosayebi, P. Šulc, J. S. Schreck, F. Romano, T. Ouldridge, R. Tsukanov, E. Nir, and A. A. Louis, "Introducing improved structural properties and salt dependence into a coarse-grained model of DNA," J. Chem. Phys. **142**, 234901 (2015).

[33] M. C. Engel, D. M. Smith, M. A. Jobst, M. Sajfutdinow, T. Liedl, F. Romano, L. Rovigatti, A. A. Louis, and J. P. K. Doye, "Force-Induced Unravelling of DNA Origami," ACS Nano **12**, 6734 (2018).

[34] C. Matek, T. E. Ouldridge, J. P. K. Doye, and A. A. Louis, "Plectoneme tip bubbles: coupled denaturation and writhing in supercoiled DNA," Sci. Rep. **5**, 7655 (2015).

[35] Z. Bryant, M. Stone, J. Gore, S. Smith, N. Cozzarelli, and C. Bustamante, "Structural transitions and elasticity from torque measurements on DNA," Nature **424**, 338–341 (2003).

[36] J. M. Schurr, "A possible cooperative structural transition of DNA in the 0.25 to 2.0 pN range," J. Phys. Chem. B **119**, 6389–6400 (2015).

[37] A. Perez, F. Lankas, F. J. Luque, and M. Orozco, "Towards a molecular dynamics consensus view of B-DNA flexibility," Nucl. Acids Res. **36**, 2379–2394 (2008).