

Review

Theoretical Methods for Studying DNA Structural Transitions under Applied Mechanical Constraints

Artem K. Efremov ^{1,2,*}, Rickson S. Winardhi ^{3,*} and Jie Yan ^{1,2,3,*}¹ Mechanobiology Institute, National University of Singapore, Singapore 117411, Singapore² Centre for Bioimaging Sciences, National University of Singapore, Singapore 117557, Singapore³ Department of Physics, National University of Singapore, Singapore 117551, Singapore

* Correspondence: mbiay@nus.edu.sg (A.K.E.); phyrsw@nus.edu.sg (R.S.W.); phyyj@nus.edu.sg (J.Y.); Tel.: +65-6516-2620 (J.Y.)

Academic Editor: Martin Kröger

Received: 6 January 2017; Accepted: 14 February 2017; Published: 21 February 2017

Abstract: Recent progress in single-molecule manipulation technologies has made it possible to exert force and torque on individual DNA biopolymers to probe their mechanical stability and interaction with various DNA-binding proteins. It was revealed in these experiments that the DNA structure and formation of nucleoprotein complexes by DNA-architectural proteins can be strongly modulated by an intricate interplay between the entropic elasticity of DNA and its global topology, which is closely related to the mechanical constraints applied to the DNA. Detailed understanding of the physical processes underlying the DNA behavior observed in single-molecule experiments requires the development of a general theoretical framework, which turned out to be a rather challenging task. Here, we review recent advances in theoretical methods that can be used to interpret single-molecule manipulation experiments on DNA.

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

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

1. Introduction

Deoxyribonucleic acid (DNA) is a fundamental biomolecule playing the major role in the preservation of genetic information across subsequent generations of living organisms. In nature, DNA can exist in several alternative forms, with the most important one being the DNA duplex consisting of two single-stranded polymers that are wound around each other forming a double helix structure. Both of the helix strands are made of a chain of basic building blocks, each containing a phosphate group, a deoxyribose sugar and one of four distinct nitrogenous bases: adenine (A), thymine (T), guanine (G) and cytosine (C). The architecture of the DNA duplex in general depends on the relative orientations of the bases, determining the way the two DNA strands interact with each other. At physiological conditions, DNA preferentially stays in the B-form, which is a right-handed duplex with a helical repeat of $h_B = 10.4$ base-pairs (bp) and a base-pair rise of $\rho_B = 0.33$ nm [1,2]. Highly specific formation of hydrogen bonds between A:T and G:C base-pairs 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 B-DNA.

Besides B-form, the DNA duplex can exist in several other states, including left-handed Z- and L-DNA with helical repeats of $h_Z \sim 12$ bp [3] and $h_L \sim 16$ bp [4–6], elongated right-handed underwound S-DNA ($h_S \sim 35$ bp [4,7–9]) and overwound P-DNA ($h_P \sim 3$ bp [4,9–12]). Each of these DNA structures has its own architecture, leading to very distinct mechanical properties. While the detailed molecular mechanisms causing DNA transition from one form to another are

not completely understood, the existing experimental data suggest that the relative stability of these structures depends on several biochemical factors, including DNA nucleotide sequence [5,6,13–17] and surrounding environmental conditions, such as solution ionic strength [13,17–19] and the level of DNA hydration [20]. More importantly, numerous single-DNA manipulation experiments suggest that the DNA transition between alternative structural states can be strongly modulated by mechanical constraints applied to the DNA duplex [4–11,16,18,19,21–27].

Indeed, it has been shown in *in vitro* experiments 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 [4–6,9–11,21–23,28]. In the former case, DNA typically develops plectonemes, structures in which DNA is twisted into a helical braid-like conformation [29,30]; whereas, in the latter scenario, DNA experiences transition from B- to L- and P-DNA forms. While L-DNA is a left-handed duplex that forms when DNA is placed under sufficiently large negative torsional stress ($\lesssim -10$ pN·nm) [4–6], at high positive torques ($\gtrsim 35$ – 40 pN·nm), DNA switches into a right-handed P-DNA state [4,11]. Furthermore, single-DNA manipulation studies show that if the force applied to a torsionally-relaxed DNA exceeds ~ 65 pN, DNA undergoes overstretching transition from B- to S-DNA state [7,8]. Therefore, DNA conformation can be subject to various mechanical constraints, which may take place in living cells.

Existing experimental studies demonstrated that DNA packaging into a highly ordered chromatin structure inside living cells, as well as translation of the genetic information coded by DNA, are performed by numerous DNA-binding proteins, which frequently distort the DNA structure by inducing mechanical deformations of the DNA duplex [31–43]. Furthermore, an increasing body of evidence has shown that not only cells impose mechanical constraints on the chromosomal DNA at normal physiological conditions to control its global conformation, but more importantly, it has been revealed that living cells can actually sense and process various forces and torques exerted on the DNA to make critical decisions. For example, several recent experimental studies have demonstrated that mechanical forces applied to the nucleus may result in transcriptional activation of specific genes [44–47]. However, the molecular mechanisms by which living cells sense the applied force and respond to it by regulating the respective intracellular processes remain unclear.

To better comprehend the physiological functions of various mechanical constraints exerted on the chromosomal DNA that may exist in living cells, it is necessary to understand how the structure of DNA and its interaction with DNA-binding proteins are regulated by different mechanical stimuli. Current biochemical bulk methods are not suited to address such questions since they do not allow the application of mechanical constraints to DNA. On the other hand, recently-developed single-DNA manipulation techniques have made it possible to directly assess the DNA interaction with various proteins under different force and torque constraints. However, the molecular mechanisms responsible for the modulation of protein binding to DNA in response to these constraints are still not fully understood. To obtain more detailed insights into the underlying physical processes, the development of a general theoretical framework describing the DNA structural stability and DNA-protein interaction is required to interpret the experimental results obtained in single-DNA manipulation experiments. In this article, we analyze recent progress on theoretical approaches aimed at the construction of such a framework and discuss possible practical applications of the reviewed methods.

2. Conformational Energy of DNA

In order to describe the DNA behavior under mechanical constraints, it is important to know how the total energy of DNA depends on its global conformation. Numerous existing theoretical and experimental studies suggest that the mechanical properties of DNA in the general case can be very well approximated by the thin rod elastic theory [30,48–53]. Under this approximation, the bending, E_{bend} , and twisting, E_{twist} , elastic energies of a short DNA fragment of length b are: $E_{\text{bend}} = \frac{L_p}{2b} \theta^2$ and

$E_{\text{twist}} = \frac{L_{\text{tw}}}{2b} \varphi^2$. Here and below, all energies are presented in $k_B T$ units, where k_B is the Boltzmann constant and T is the temperature of the surrounding environment. In the above formulas, θ and φ are the bending and twisting angles of the short DNA fragment, respectively. As for the parameters L_p and L_{tw} , which have a dimension of length (due to the energy normalization by $k_B T$), they are often referred to as the bending and twisting persistence lengths of DNA. These two parameters describe the bending and twisting rigidities of DNA; the larger the values of L_p and L_{tw} , the stiffer the DNA is. Finally, the elastic stretching deformation of DNA is typically neglected in theoretical studies, as it has been previously shown that even under very high tensions (50–60 pN), the relative change of the DNA contour length is <5% [54,55], which contributes to less than 5% of the conformational energy of the DNA. Therefore, the theoretical framework discussed in this review is mainly based on the inextensible semi-flexible polymer model of DNA.

It should be noted that the above approximations of the DNA bending and twisting elastic energies work well only when $b \ll L_p$ and $b \ll L_{\text{tw}}$, so that the thermally-excited bend and twist of the short DNA fragment are sufficiently small (i.e., $\theta \ll 1$ and $\varphi \ll 1$). In this case, it is not hard to show that $\hat{t} \cdot \hat{t}' = \cos \theta \approx 1 - \frac{1}{2}\theta^2$, where \hat{t} and \hat{t}' are the unit vectors tangent to the DNA fragment at its ends. With this small note, the bending energy of the DNA fragment can be re-written in a more convenient form as $E_{\text{bend}} = \frac{L_p}{2b} (\hat{t} - \hat{t}')^2$.

Using the above formulas for short DNA fragments, one can easily derive an expression for the conformational energy, E_{conf} , of an arbitrarily long DNA. Indeed, by dividing the DNA polymer into small straight segments of equal length b , which form a polygonal chain representation of the DNA (see Figure 1a), the conformational energy can be simply found as the sum of local elastic bending and twisting energies:

$$E_{\text{conf}} = \sum_{n=1}^{N-1} \frac{L_p}{2b} (\hat{t}_{n+1} - \hat{t}_n)^2 + \sum_{n=1}^N \frac{L_{\text{tw}}}{2b} \varphi_n^2 \tag{1}$$

where N is the total number of DNA segments in the polygonal chain representing DNA; \hat{t}_n and φ_n are the tangent vector and the twisting angle of the n -th DNA fragment, respectively.

In the continuous limit, $b \rightarrow 0$, the above sums transform into two very well-known integrals from the elastic theory of thin rods:

$$E_{\text{conf}} = \frac{L_p}{2} \int_0^L \left(\frac{d\hat{t}(s)}{ds} \right)^2 ds + \frac{L_{\text{tw}}}{2} \int_0^L \left(\frac{d\varphi(s)}{ds} \right)^2 ds \tag{2}$$

Here, L is the DNA contour length; $\hat{t}(s)$ and $\varphi(s)$ are the tangent vector and the twist of the DNA polymer at the point situated on the DNA contour that corresponds to the arc length s .

While Equations (1) and (2) describe the elastic energy of unconstrained DNA, for practical use, it is frequently important to know how DNA behaves under force, \hat{f} , and torque, τ . The application of these mechanical constraints to DNA results in the appearance of two additional energy terms, $-\hat{f} \cdot \hat{d}$ and $-2\pi\tau\Delta Lk$, in the expression for the total conformational energy of the DNA:

$$E_{\text{tot}} = E_{\text{conf}} - \hat{f} \cdot \hat{d} - 2\pi\tau\Delta Lk \tag{3}$$

Here, \hat{d} is the DNA end-to-end distance vector (i.e., $\hat{d} = b \sum_{n=1}^N \hat{t}_n$) and ΔLk is the DNA linking number change with respect to the torsionally-relaxed DNA state (see more details in Section 4). In this formula, the force \hat{f} and torque τ are scaled by $k_B T$; thus, \hat{f} has a dimension of length⁻¹, and τ is dimensionless.

Knowing the total energy, it is possible to predict the DNA behavior in the thermodynamic equilibrium since it is completely determined by the Boltzmann distribution, $\frac{1}{Z} e^{-E_{\text{tot}}}$. However, calculation of the DNA partition function, Z , in the general case can be extremely difficult. In the following sections, we review several theoretical approaches that have been developed to predict the equilibrium DNA behavior under force and torque constraints and discuss their potential applications.

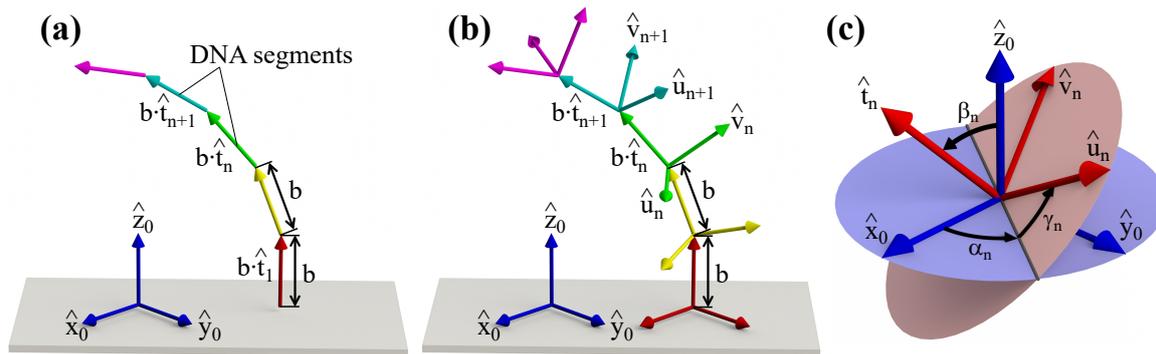


Figure 1. Discretized model of DNA. In theoretical computations, the DNA polymer is usually represented by a polygonal chain comprised of short DNA segments of equal length, b . While in the case of torsion-free DNA, the global DNA conformation is completely specified by the set of unit tangent vectors $(\hat{t}_1, \dots, \hat{t}_N)$ of the respective DNA segments (a). In a more general scenario, one needs to know the 3D-orientations of all of the DNA segments, which are treated as solid rigid bodies with attached local Cartesian coordinate frames $(\hat{u}_n, \hat{v}_n, \hat{w}_n)$ (b). (c) Schematic picture of the three Euler rotation angles $(\alpha_n, \beta_n, \gamma_n)$ determining the orientation of the n -th DNA segment with respect to the fixed global coordinate system $(\hat{x}_0, \hat{y}_0, \hat{z}_0)$.

3. Mechanically-Stretched Homogeneous DNA

In the simplest case of mechanically-stretched torsion-free homogeneous DNA, the above equations reduce to the so-called worm-like chain polymer model (WLC), and the total energy of DNA becomes [48]:

$$E_{\text{tot}} = \frac{L_p}{2} \int_0^L \left(\frac{d\hat{f}(s)}{ds} \right)^2 ds - \hat{f} \cdot \hat{d} \tag{4}$$

To quantify the net average effect of the applied tension on the DNA behavior, the force-extension curve of DNA, $z(f) = \langle \hat{d} \cdot \hat{z}_0 \rangle$, is typically measured in single-molecule experiments [54,56]. Here, z is the average DNA extension along the force direction described by the unit vector \hat{z}_0 ; $f = |\hat{f}| = \hat{f} \cdot \hat{z}_0$ is the force magnitude; and the averaging $\langle \dots \rangle$ is performed over all of the equilibrium DNA conformations weighted by the corresponding Boltzmann factor, $\frac{1}{Z} e^{-E_{\text{tot}}}$.

Detailed analysis of the above formula shows that the bending persistence length, L_p , separates the force-response of DNA into a low ($f \ll \frac{1}{L_p}$) and a high ($f \gg \frac{1}{L_p}$) force regime. In the low force regime, the DNA force-extension curve can be well approximated by the freely-jointed chain polymer (FJC) model: $L_p f = \frac{3}{2} \frac{z}{L}$; whereas in the high force regime, it exhibits highly non-linear behavior: $L_p f = [4(1 - \frac{z}{L})^2]^{-1}$ [48,49]. A simple interpolation between the two regimes leads to the famous Marko–Siggia’s formula that has been shown to be able to describe the DNA behavior over the entire force range [48]:

$$f = \frac{1}{L_p} \left[\frac{z}{L} + \frac{1}{4(1 - \frac{z}{L})^2} - \frac{1}{4} \right] \tag{5}$$

The simple yet very useful Marko–Siggia formula that relates the DNA extension change to the applied force has found many applications in the analysis of single-molecule experimental data over the ensuing years. However, the requirement of the DNA homogeneity makes it difficult to use this formula in situations where DNA inhomogeneities naturally arise as a result of sequence-dependent properties of DNA [6,57–61] or due to proteins binding to various DNA sites. The solution to this problem requires the construction of new approaches capable of taking into account local variations in the DNA structure.

4. Homogeneous DNA Subjected to Both Force and Torque Constraints

In the case of DNA subjected to both force and torque constraints, the total energy of DNA becomes:

$$E_{\text{tot}} = \frac{L_p}{2} \int_0^L \left(\frac{d\hat{t}(s)}{ds} \right)^2 ds + \frac{L_{\text{tw}}}{2} \int_0^L \left(2\pi \frac{d\text{Tw}(s)}{ds} \right)^2 ds - \hat{f} \cdot \hat{d} - 2\pi\tau\Delta\text{Lk} \tag{6}$$

Here, $\text{Tw}(s) = \frac{1}{2\pi}\varphi(s)$ is the DNA twist number at the point corresponding to the arc length s , and ΔLk is the DNA linking number change with respect to the torsionally-relaxed DNA reference state.

By definition, the linking number, Lk , represents the number of times the two strands of DNA are twisted around each other [62–67]. Thus, knowing the architecture of the DNA duplex, it is easy to find the linking number of any torsion-free DNA structure as the ratio of the total DNA length to the helical repeat of the respective DNA form. For example, the helical repeat of DNA in B-form is $h_B = 10.4$ base-pairs per turn. Hence, the linking number of a torsion-free B-DNA comprised of N_{bp} base-pairs is: $\text{Lk}_{0,B} = N_{\text{bp}}/h_B$.

Since DNA predominantly stays in B-form under physiological conditions, it is typically used as a reference state for all of the calculations based on Equation (6) with the DNA linking number change defined as: $\Delta\text{Lk} = \text{Lk} - \text{Lk}_{0,B}$. It should be noted that at a fixed force, \hat{f} , and torque, τ , applied to DNA, its average linking number $\langle \text{Lk} \rangle$ is simply proportional to the DNA length. Due to this reason, it is frequently more convenient to express the level of DNA supercoiling in terms of the DNA length-independent quantity $\sigma = \Delta\text{Lk}/\text{Lk}_{0,B}$, which is usually referred to as the superhelical density/linking number density/specific linking difference. Negative/positive value of σ corresponds to negatively (underwound) or positively (overwound) supercoiled DNA, respectively. This number shows how many turns are removed from/added to the DNA per a single turn of the two strands in the relaxed configuration.

Using Călugăreanu–White’s theorem [62,63], the DNA linking number change, ΔLk , can be conveniently expressed as a sum of two components: $\Delta\text{Lk} = \text{Tw} + \text{Wr}$, where Tw and Wr are the DNA twist and writhe numbers:

$$\begin{aligned} \text{Tw} &= \int_0^L \frac{d\text{Tw}(s)}{ds} ds = \frac{1}{2\pi} \int_0^L \left[\hat{u}(s) \times \frac{d\hat{u}(s)}{ds} \right] \cdot \hat{t}(s) ds \\ \text{Wr} &= \frac{1}{4\pi} \int_0^L \int_0^L [\hat{t}(s) \times \hat{t}(s')] \cdot \frac{\hat{r}(s) - \hat{r}(s')}{|\hat{r}(s) - \hat{r}(s')|^3} ds ds' \end{aligned} \tag{7}$$

Here, $\hat{r}(s)$ and $\hat{r}(s')$ are the position vectors of points situated on the DNA polymer that correspond to the arc lengths s and s' , respectively; $\hat{t}(s)$ is the unit tangent vector, and $\hat{u}(s)$ is the unit vector normal to the DNA contour at the point s , such that $\hat{u}(s)$ keeps track of the DNA twist, $\text{Tw}(s)$.

Due to the double-contour integral in the above expression for the DNA writhe number, it is clear that in general, Wr depends on the global DNA conformation. However, it is usually more convenient to use a simpler Fuller’s formula [67], which provides accurate estimates of the DNA writhe number for certain cases described below:

$$\text{Wr}^F = \frac{1}{2\pi} \int_0^L \frac{[\hat{t}(s) \times \frac{d\hat{t}(s)}{ds}] \cdot \hat{z}_0}{1 + \hat{t}(s) \cdot \hat{z}_0} ds \tag{8}$$

where, as before, \hat{z}_0 is the unit vector pointing in the direction of force \hat{f} applied to the DNA. It can be shown that $\text{Wr}^F = \text{Wr}$ for those DNA conformations that can be obtained by a continuous deformation of DNA initially extended along the \hat{z}_0 -axis direction in such a way that none of the intermediate DNA conformations are represented by a self-intersecting curve or a curve that has points at which the denominator of the integrand in Equation (8) resets to zero [50,67]. For example,

Fuller’s formula accurately describes the writhe number of nearly extended or solenoid-shaped DNA conformations [51,52]. Thus, it can be used to predict the behavior of DNA under a wide range of mechanical constraints up to the onset of the torque-induced buckling transition when DNA starts to develop plectonemes [50–53,68,69].

While Equations (6) and (7) provide a general analytical background for theoretical studies of supercoiled DNA, in real calculations, it is often more convenient to use discretized versions of these formulas where DNA is modeled as a polygonal chain consisting of short segments of equal length b ($b \ll L_p$ and $b \ll L_{tw}$); see Figure 1b. The orientation of each DNA segment in such a model is represented by a local Cartesian coordinate frame $(\hat{u}_n, \hat{v}_n, \hat{t}_n)$ obtained by a three-dimensional rotation about the origin of a fixed lab coordinate system $(\hat{x}_0, \hat{y}_0, \hat{z}_0)$: $\hat{u}_n = R_n \hat{x}_0$, $\hat{v}_n = R_n \hat{y}_0$, and $\hat{t}_n = R_n \hat{z}_0$. Here, n is the index of the corresponding DNA segment, and $R_n = R_{\alpha_n} R_{\beta_n} R_{\gamma_n}$ denotes a rotation matrix resulting from the composition of three successive Euler revolutions through angles α_n , β_n and γ_n shown in Figure 1c (where $\alpha_n, \gamma_n \in [0, 2\pi]$ and $\beta_n \in [0, \pi]$) [69]. Since the Cartesian coordinate frame of each DNA segment $(\hat{u}_n, \hat{v}_n, \hat{t}_n)$ is in a one-to-one relation with the respective rotational matrix R_n , hereafter, we simply use R_n to designate the n -th DNA segment orientation.

Using the polygonal chain representation of DNA, it can be shown that Equation (7) can be re-written in the following discretized form [70,71]:

$$\begin{aligned} Tw &= \sum_{n=1}^{N-1} \Delta Tw_n(R_n, R_{n+1}) \\ Wr &= \frac{1}{4\pi} \sum_{\substack{n, n'=1 \\ n \neq n'}}^N \Omega_{nn'} \end{aligned} \tag{9}$$

where $\Delta Tw_n(R_n, R_{n+1}) = \frac{1}{2\pi} [\hat{u}_n \times \hat{u}_{n+1}] \cdot \hat{t}_n = \frac{1}{2\pi} [R_n \hat{x}_0 \times R_{n+1} \hat{x}_0] \cdot R_n \hat{z}_0$ is the twist number corresponding to the local contribution of the n -th and $(n + 1)$ -th DNA segments, which is simply the twist angle, φ_n , between the coordinate systems described by rotation matrices R_n and R_{n+1} normalized to 2π ; and $\Omega_{nn'}$ is the solid angle corresponding to the quadrangle formed by the n -th and n' -th DNA segments (see Figure 2), which can be found by using Equations (14)–(16) from [71].

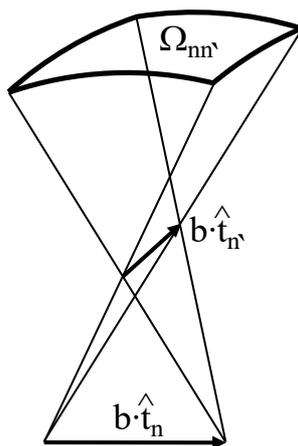


Figure 2. Solid angle $\Omega_{nn'}$ corresponding to the quadrangle formed by the n -th and n' -th DNA segments.

Combining together Equations (1), (3) and (9) and taking into account that $\hat{f} \cdot \hat{d} = \hat{f} \cdot b \sum_{n=1}^N \hat{t}_n = bf \sum_{n=1}^N \hat{z}_0 \cdot R_n \hat{z}_0$, it is straightforward to obtain the discretized version for the total conformational energy of DNA:

$$E_{\text{tot}}(R_1, \dots, R_N) = \sum_{n=1}^{N-1} \frac{L_p}{2b} [R_{n+1} \hat{z}_0 - R_n \hat{z}_0]^2 + \sum_{n=1}^{N-1} \frac{L_{\text{tw}}}{2b} [2\pi \Delta T w_n(R_n, R_{n+1})]^2 - bf \sum_{n=1}^N \hat{z}_0 \cdot R_n \hat{z}_0 - 2\pi\tau \left[\sum_{n=1}^{N-1} \Delta T w_n(R_n, R_{n+1}) + \frac{1}{4\pi} \sum_{\substack{n,n'=1 \\ n \neq n'}}^N \Omega_{nn'} \right] \quad (10)$$

5. Metropolis-Monte Carlo Simulations

Having at hand the formula for the total conformational energy of DNA (Equation (10)), one can apply the Metropolis-Monte Carlo algorithm [72,73] to computationally generate equilibrium DNA conformations that satisfy the Boltzmann distribution [74,75]. By getting a sufficiently large number of such DNA states, it is then possible to predict how various observable DNA characteristics, including the average DNA extension and linking number change, evolve in response to mechanical constraints applied to the DNA.

Figure 3 shows examples of representative polymer conformations obtained via the Metropolis-Monte Carlo algorithm for mechanically-stretched DNA at two different torques. As can be seen from the figure, in the case of torsionally-relaxed DNA ($\tau = 0$), the polymer typically assumes an extended conformation, which does not show preference for the formation of either positively- or negatively-coiled DNA structures; see Figure 3a. Upon the increase of the torque magnitude, the DNA undergoes the so-called buckling transition, as shown in Figure 3b, featured by the formation of multiple DNA writhes, which eventually develop into supercoiled DNA plectonemes; see Figure 3c.

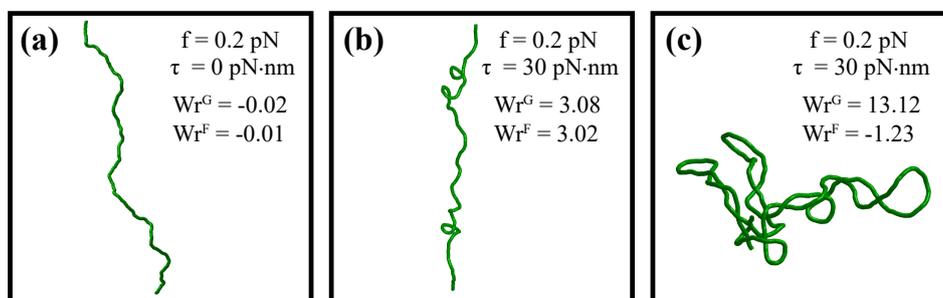


Figure 3. Examples of DNA equilibrium conformations obtained by the Metropolis-Monte Carlo algorithm. In the absence of torque, mechanically-stretched DNA typically assumes an extended chain configuration (a). On the other hand, the application of sufficiently large positive or negative torque leads to the DNA buckling transition (b), which eventually results in the formation of supercoiled DNA plectoneme structures (c). Evaluation of the DNA writhe number by the Gauss double contour integral (Wr^G) and Fuller's formula (Wr^F) shows that both formulas predict very similar values of the DNA writhe number up to the onset of the DNA buckling transition. For plectonemic DNA structures, however, Fuller's formula does not provide accurate estimations. In the above simulations, the length of DNA was set to 0.85 μm .

It is interesting to note that the estimations of the DNA writhe number by the Gauss double contour integral, Wr^G (Equation (7)) and Fuller's formula, Wr^F (Equation (8)) agree with each other very well up to the onset of the DNA buckling transition, in full accordance with the previously published theoretical studies [50–53]. However, as soon as DNA starts to form supercoiled plectonemes, the writhe number predicted by Fuller's formula starts to deviate from the one computed from the

Gauss double contour integral. Thus, any interpretations of the observed DNA behavior involving Fuller's formula, which is frequently used in theoretical studies of DNA, should be restricted up to the onset of the DNA buckling transition [50,51].

Metropolis-Monte Carlo simulations have been widely used to investigate the response of DNA to various mechanical constraints, providing valuable quantitative description and interpretation of numerous single-DNA manipulation experiments [52,53,74–76]. However, in order to obtain sufficiently large and accurate Boltzmann statistics of equilibrium DNA states, a considerable number of the algorithm iterations have to be carried out, which typically requires extensive computational resources. Together with the lack of detailed analytical insights into the DNA behavior, this imposes significant restrictions on the predictive power of such a numerical approach.

6. Transfer-Matrix Formalism

6.1. Transfer-Matrix Calculations of the DNA Partition Function

In order to gain more detailed understanding of the DNA micromechanics under force and torque constraints, several analytic polymer theories have been developed during recent years [12,50–53,76–78]. While these studies resulted in many important insights into the equilibrium behavior of DNA, most of them were mainly based on the assumption of the uniformity of the DNA physical properties without taking into account possible local variations, which may either come from the sequence-dependence of the DNA parameters, local DNA transitions into alternative structural states or due to proteins binding to various DNA sites.

In contrast to the above theories, the recently-described transfer-matrix formalism allows one to take into consideration both local deformations and structural transformations of DNA [68,69,79–81]. The main idea of the method is based on the nearest neighbor approximation. It has been shown that if the total energy of DNA can be written as the sum of local DNA segments contributions, $E_{\text{tot}} = \sum_{n=1}^{N-1} E_n(R_n, R_{n+1})$, then the DNA partition function, Z , can be found as a trace of the ordered product of transfer-matrices S_n defined on the vertices that join neighboring segments in the discretized DNA model: $Z \sim \text{Tr}(\prod_{n=1}^{N-1} S_n)$; see [68,69,79,80] and the details below. Here, $E_n(R_n, R_{n+1})$ is the energy contribution corresponding to the vertex connecting the n -th and $(n + 1)$ -th segments of the polygonal chain representing the DNA polymer. As matrices S_n depend only on the local physical properties of the respective DNA segments, it is clear that the transfer-matrix calculations do not require the assumption of the DNA homogeneity.

Because of these advantages, the transfer-matrix approach may provide a wide scope of potential applications, such as the investigation of sequence-dependent DNA response to mechanical constraints and the exploration of the interactions between DNA and architectural proteins. However, the requirement of the DNA total conformational energy being expressed as the sum of local DNA segments contributions imposes certain limits on the transfer-matrix method applications in the case of DNA subjected to torque constraints.

Indeed, as was mentioned in Section 4, the DNA writhe number generally depends on the global DNA conformation and cannot be written as a simple sum of local DNA segments' contributions; see Equation (9). This limitation can be partially overcome by utilizing Fuller's formula (Equation (8)), which provides correct estimates of the DNA writhe number up to the onset of the torque-induced buckling transition when DNA starts to develop supercoiled plectonemes; see [51–53] and Figure 3. Although the application of Fuller's formula beyond the DNA buckling transition point results in inaccurate evaluation of the DNA writhe number (see Figure 3c), it has been shown in our recent work that with a slight revision of the expression for the DNA total conformational energy, the transfer-matrix formalism still can be used to predict the DNA behavior in a wide range of force and torque constraints [69]. Namely, by introducing an additional energy term, $\tau \lambda W r^F$, with a free model parameter, λ , it is possible to describe the force- and torque-response of DNA up to

the experimentally-measured DNA buckling transition boundary, which can be reproduced in the transfer-matrix calculations simply by tuning the value of λ .

Taking into account the above notes and introducing the new energy term $\tau\lambda Wr^F = \tau\lambda(\Delta Lk^F - Tw)$ into Equation (10), it can be shown that the total conformational energy of DNA can be expressed with the help of Fuller’s formula as [69]:

$$E_{tot}(R_1, \dots, R_N) = \sum_{n=1}^{N-1} E_n(R_n, R_{n+1}) - bf(\hat{z}_0 \cdot R_N \hat{z}_0) \tag{11}$$

where the local energy terms corresponding to neighboring DNA segments contributions assume the following form:

$$E_n(R_n, R_{n+1}) = \frac{a}{2} (R_n \hat{z}_0 - R_{n+1} \hat{z}_0)^2 + \frac{c}{2} (2\pi\Delta Tw_n)^2 - bf(\hat{z}_0 \cdot R_n \hat{z}_0) - \tau(2\pi - \lambda)\Delta Lk_n^F - \tau\lambda\Delta Tw_n \tag{12}$$

Here, $a = L_p/b$ and $c = L_{tw}/b$ are dimensionless parameters describing the bending and twisting rigidities of DNA segments in the discretized model; $\Delta Tw_n = \Delta Tw_n(R_n, R_{n+1})$ and $\Delta Lk_n^F = \frac{1}{2\pi}(\tilde{\alpha}_{n+1} + \tilde{\gamma}_{n+1} - \tilde{\alpha}_n - \tilde{\gamma}_n)$ are local DNA segments contributions to the total DNA twist and linking number change, respectively [50,69], where $\tilde{\alpha}_n$ and $\tilde{\gamma}_n$ are the Euler angles of the n -th DNA segment from the extended range of $(-\infty, +\infty)$.

Knowing the total conformational energy of DNA under tension, f , and torque, τ , its partition function, $Z_{f,\tau}$, can be found as:

$$Z_{f,\tau} = \int dR_1 \dots dR_N e^{-E_{tot}(R_1, \dots, R_N)} \zeta(R_N, R_1) = \int dR_1 \dots dR_N \prod_{n=1}^{N-1} T(R_n, R_{n+1}) \times \sigma(R_N, R_1) \tag{13}$$

where $\zeta(R_N, R_1)$ is a function that imposes specific boundary conditions on the orientations of the DNA ends; $\sigma(R_N, R_1) = \zeta(R_N, R_1)e^{bf(\hat{z}_0 \cdot R_N \hat{z}_0)}$ and $T(R_n, R_{n+1}) = e^{-E_n(R_n, R_{n+1})}$. In the above formula, the integration is carried out over all of the possible DNA segments’ orientations, i.e., $\int dR_n = \int_0^{2\pi} d\alpha_n \int_0^{2\pi} d\gamma_n \int_0^\pi \sin \beta_n d\beta_n$.

Equation (13) can be further simplified by recalling that any square-integrable function, ψ , defined on the SO(3) group of 3D rotation matrices parametrized by Euler angles α, β, γ can be expanded into a series of orthogonal D-functions, $D_{ml}^n(\alpha, \beta, \gamma)$ [82]. The expansion coefficients of this series can be found as $\psi_{mln} = \frac{2n+1}{8\pi^2} \int dR \psi(R) \overline{D_{ml}^n}(R)$, where for the sake of simplicity, we use the following notations: $\psi(R) = \psi(\alpha, \beta, \gamma)$ and $D_{ml}^n(R) = D_{ml}^n(\alpha, \beta, \gamma)$. The bar over function D_{ml}^n in the above formula denotes complex conjugation. Performing such an expansion for functions $\sigma(R_N, R_1)$ and $T(R_n, R_{n+1})$, it can be shown that for a sufficiently long DNA ($L \gg L_p$ and $L \gg L_{tw}$), Equation (13) reduces to a mere multiplication of matrices whose entries are the expansion coefficients of $T(R_n, R_{n+1})$ and $\sigma(R_N, R_1)$ functions (see [69] for more details):

$$Z_{f,\tau} = \text{Tr}(S^{N-1}V) \tag{14}$$

where the elements of matrices S and V are:

$$\begin{aligned} S_{kk'}(a, b, c, \lambda) &= \pi^2 \sqrt{(2k+1)(2k'+1)} e^{-a-c} \\ &\times \sum_{p,s,s',r} (2p+1)(2s+1)(2s'+1) e^{ir(\omega - \frac{\pi}{2})} I_r(\tau[1 - \frac{\lambda}{2\pi}]) I_r(c\sqrt{1+\chi^2}) \\ &\times \mathcal{L}_r^p(-a) \mathcal{L}_r^s(-bf) \mathcal{L}_r^{s'}(0) \begin{pmatrix} p & s & k \\ -r & r & 0 \end{pmatrix}^2 \begin{pmatrix} p & s' & k' \\ -r & r & 0 \end{pmatrix}^2 \end{aligned} \tag{15}$$

and:

$$V_{kk'}(b) = 8\pi^2 \delta_{k'0} i_k(bf) \sqrt{2k+1} \tag{16}$$

Here, i is the imaginary unit; $\chi = \frac{\tau\lambda}{2\pi c}$ and $\omega = \tan^{-1}(\chi)$; $I_r(x)$ are modified Bessel functions of the first kind; $i_k(x)$ are spherical modified Bessel functions of the first kind; $\mathcal{L}_r^k(s) = \int_{-1}^1 d_{rr}^k(\cos^{-1} x) e^{-sx} dx$ is the bilateral Laplace transform of the diagonal elements, d_{rr}^k , of the Wigner small d-matrix; and $\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}$ are the Wigner 3-j symbols. Note that the parallel boundary condition ($\hat{t}_1 = \hat{t}_N = \hat{z}_0$), which was originally used to derive Equations (15) and (16), does not affect the final results as long as the DNA contour length is much larger than the DNA bending and twisting persistence lengths ($L \gg L_p$ and $L \gg L_{tw}$) [69].

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 $\Delta Lk_{f_0}(\tau) = \frac{1}{2\pi} \frac{\partial \ln(Z_{f=f_0,\tau})}{\partial \tau} |_{\chi=\text{const}}$.

Combining these formulas with Equations (14)–(16), it is possible to predict the DNA behavior in a wide range of mechanical constraints, assuming that DNA does not experience structural transitions between alternative forms [69]. However, as has been mentioned in the Introduction (Section 1), numerous experimental data show that DNA frequently undergoes structural transformations when it is subjected to sufficiently large force and/or torque. In the next two sections, we review the application of the transfer-matrix formalism to understanding the force- and torque-dependent stability of such DNA structures.

6.2. Description of the DNA Structural Transitions Using the Transfer-Matrix Formalism

From single-molecule experiments, it is known that double-stranded DNA can exist in several structural states, including B-, L-, P- and S-DNA, which have very different geometric characteristics and material properties [2,4–6,8–11,16,21–23,28,54,56]. Furthermore, it has been found that the relative stability of these structures sensitively depends on the mechanical constraints applied to DNA. Understanding of the physical mechanisms responsible for the observed force- and torque-induced DNA transitions between various DNA forms has imposed a huge theoretical challenge to the field.

Several alternative approaches have been developed so far to solve the above problem [12,69,76,78] with two of the most recent methods being based on the transfer-matrix calculations and the evaluation of the energy global minima state of various DNA structures [12,69]. While the latter approach rests on the assumption of DNA homogeneity similar to many previous theories, the transfer-matrix method is free from such limitation and, in fact, requires only minor changes in the above formulas in order to incorporate various DNA structures into the partition function computations.

Indeed, by taking B-DNA as the reference state, transitions from B- to L-, P- or S-DNA forms are determined by the energy difference between the respective DNA structures, which includes a contribution from the base-pairing energy difference, μ_u , and the respective contributions from the applied force, $-b_u f(\hat{z}_0 \cdot R_n \hat{z}_0)$, and torque, $-2\pi\tau \Delta l k_0^{(u)}$. Here, $u = L, P, S$ or B denotes the structural state of the DNA segments in the discretized model of DNA; b_u is the length of DNA segments in the respective state, u ; and $\Delta l k_0^{(u)} = l k_{0,u} - l k_{0,B}$, where $l k_{0,u} = \pm h_u^{-1}$ is the relaxed linking number of the respective DNA form per single base-pair, which is assigned to be positive for right-handed DNA helical structures (like B-, P- or S-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 $\Delta l k_0^{(B)} = 0$.

Utilizing the above notations, the total conformational energy of DNA in the general case can be found as:

$$E_{\text{tot}}(u_1, \dots, u_N, R_1, \dots, R_N) = \sum_{n=1}^{N-1} E_n^{(u_n)}(R_n, R_{n+1}) + q \sum_{n=1}^N \left(\mu_{u_n} - 2\pi\tau \Delta l k_0^{(u_n)} \right) - b_{u_N} f(\hat{z}_0 \cdot R_N \hat{z}_0) \tag{17}$$

where (R_1, \dots, R_N) and (u_1, \dots, u_N) are the orientations and structural states of DNA segments in the discretized model of DNA; q is the number of base-pairs in each segment; $b_u = q\rho_u$ is the length of the DNA segments in state u , where ρ_u is the base-pair rise in the corresponding DNA structure. As for the local energy terms, $E_n^{(u_n)}(R_n, R_{n+1})$, they have absolutely the same form as in Equation (12) with the only difference being that parameters $a_u = L_p^{(u)}/b_u$, $c_u = L_{tw}^{(u)}/b_u$ and λ_u describing the respective physical properties of DNA segments now depend on the segments' state, $u = L, P, S$ or B . Here, $L_p^{(u)}$ and $L_{tw}^{(u)}$ are the bending and twisting persistence lengths of DNA in state u . The last term in Equation (17) corresponding to the potential energy of the DNA end segment is negligible in the general case compared to the total conformational energy of DNA and is kept in Equation (17) solely for the purpose of mathematical rigorousness.

Since Equations (11) and (17) have a very similar appearance, it can be shown that the transfer matrix, S_u , for DNA segments being in state u can be obtained simply via multiplication of Equation (15) by the exponential factor $e^{-q(\mu_u - 2\pi\tau\Delta lk_0^{(u)})}$ [69]:

$$(S_u)_{kk'} = e^{-q(\mu_u - 2\pi\tau\Delta lk_0^{(u)})} S_{kk'}(a_u, b_u, c_u, \lambda_u) \tag{18}$$

Then, by using Equation (18), the DNA partition function in the general case can be found as (see [69] for more details):

$$Z_{f,\tau} = \text{Tr}(US^{\hat{N}-1}\hat{V}) \tag{19}$$

where $U = \begin{pmatrix} I & I & I & I \end{pmatrix}$ is a block-matrix comprising four identity matrices, I ; and the transfer matrix, \hat{S} , and boundary condition matrix, \hat{V} , are:

$$\hat{S} = \begin{pmatrix} S_B & S_B & S_B & S_B \\ S_L & S_L & S_L & S_L \\ S_P & S_P & S_P & S_P \\ S_S & S_S & S_S & S_S \end{pmatrix}, \quad \text{and} \quad \hat{V} = \begin{pmatrix} V_B \\ e^{-q(\mu_L - 2\pi\tau\Delta lk_0^{(L)})} V_L \\ e^{-q(\mu_P - 2\pi\tau\Delta lk_0^{(P)})} V_P \\ e^{-q(\mu_S - 2\pi\tau\Delta lk_0^{(S)})} V_S \end{pmatrix} \tag{20}$$

Here, blocks S_u are defined by Equation (18), and blocks $V_u = V(b_u)$ are described by the previously introduced Equation (16).

By performing differentiation of the DNA partition function defined by Equation (19) with respect to the force, f , or torque, τ , as was described at the end of Section 6.1, it is straightforward to calculate the DNA force-extension and/or the torque-extension curves under various mechanical constraints. The results of such computations, as well as their discussions can be found in the next section.

7. Force- and Torque-Dependent DNA Structural Transitions

The above transfer-matrix approach provides a general theoretical framework for analyzing data collected in single-DNA stretching and twisting experiments. To demonstrate its application, we calculated the force-extension curves, $z_{\tau_0}(f)$, and force-superhelical density curves, $\sigma_{\tau_0}(f)$, under a constant torque ($\tau = \tau_0$), as well as torque-extension curves, $z_{f_0}(\tau)$, and torque-superhelical density curves, $\sigma_{f_0}(\tau)$, under a constant force ($f = f_0$) constraints, for DNA that can switch between different structural forms. Here, we report results extending the previously published theoretical data from [69] to a larger force range by including S-DNA structural state in the transfer-matrix computations to make the physical picture more complete.

In all of our calculations, the size of the DNA segments was set to be equal to $q = 1.5$ bp for all of the DNA structures. The rest of the parameters used in the transfer-matrix analysis are summarized in Table 1, with reference to the experimental and theoretical DNA studies from which the parameters' values have been extracted.

Table 1. Structural and elastic characteristics of various DNA forms.

DNA form, u	Bending persistence length, $L_p^{(u)}$ (nm)	Twisting persistence length, $L_{tw}^{(u)}$ (nm)	Contour length relative to B-DNA form	DNA helical repeat, h_u (bp)	Base-pairing energy relative to B-DNA form, μ_u ($k_B T$)	λ_u
B-DNA	50 [54,56]	95 [4,23,28]	1	10.4 [2]	0	4.3 [69]
L-DNA	7 [5,12]	15 [5,6,12]	1.35 [5,12]	16 [4–6]	5.0 [69]	4.3 [69]
P-DNA	15 [12]	25 [12]	1.7 [9,10,12]	3 [4,9–12]	17.8 [69]	−0.5 [69]
S-DNA	15 [12,83]	20 [12]	1.7 [4,8,9]	35 [4,9]	5.1 (this study)	4.3 (this study)

As can be seen from Figure 4a,b,d,e, the DNA force-extension and force-superhelical density curves demonstrate two types of behavior: (a) smooth continuous variation with the applied force corresponding to the force-response of a particular DNA form; and (b) discontinuous, stepwise change, which indicates DNA transition from one structural state to another.

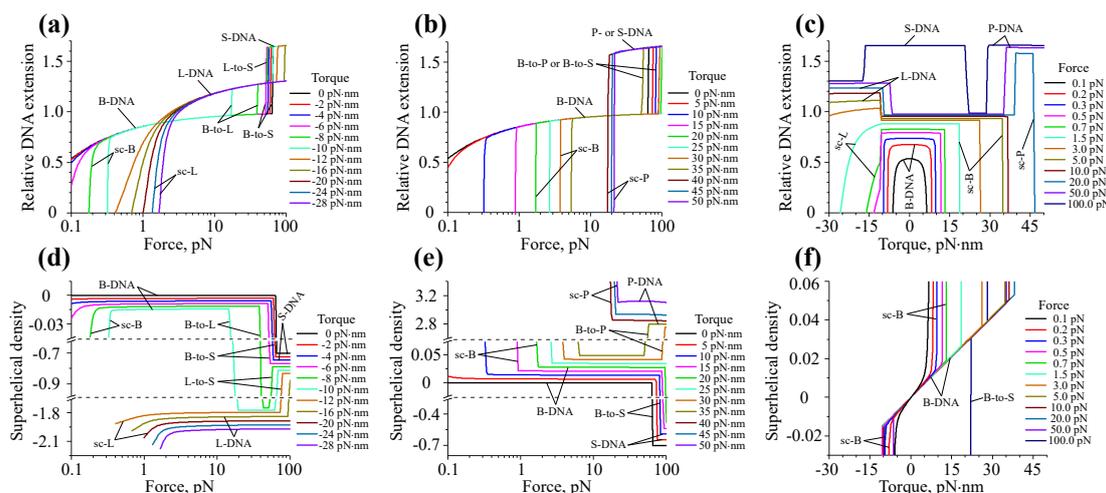


Figure 4. Structural stability of DNA under various mechanical constraints. (a,b) DNA force-extension curves, $z_{\tau_0}(f)$, and (d,e) force-superhelical density curves, $\sigma_{\tau_0}(f)$, calculated at different values of the applied torque ($\tau = \tau_0$). Panels (a,d) and (b,e) display results obtained for negative and positive torques, respectively. Abrupt changes in the behavior of the DNA force-extension curves presented on panels (a,b) mark the DNA transitions from B-form into alternative L-, P- or S-DNA structures, which are denoted by the B-to-L, B-to-P, B-to-S and L-to-S tags. The graphs on panels (d,e) show that these transitions are accompanied by a drastic change of the DNA superhelical density due to the large difference in the relaxed linking numbers of the corresponding DNA forms. (c) DNA torque-extension curves, $z_{f_0}(\tau)$, and (f) torque-superhelical density curves, $\sigma_{f_0}(\tau)$, at different values of the applied force ($f = f_0$). 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. In panels (a–c), the DNA extension is normalized to the total contour length of DNA in B-form. The following abbreviations are used to indicate different DNA states and structural transitions in the above plots: sc-B, sc-L and sc-P, supercoiled B-, L- and P-DNA, respectively; B-to-L, B-to-P, B-to-S and L-to-S, structural transitions between the respective DNA forms.

While at small torques DNA stays in the B-form and has a preferentially extended conformation, application of a moderate torsional stress to DNA results in the development of supercoiled structures (sc-B), whose formation is manifested by a rapid drop of the DNA extension and a corresponding change in the DNA superhelical density; see Figure 4a,b,d,e. A further increase of the torque magnitude leads to the B-to-L structural transition of DNA at high negative torques ($\lesssim -10$ pN·nm) and B-to-P transition at high positive torques ($\gtrsim 35$ – 40 pN·nm). Both of these DNA transformations are accompanied by a steep change in the DNA superhelical density seen on Figure 4d,e, which results from the very large difference in the relaxed linking numbers of B-, L- and

P-DNA forms. Altogether, these data suggest that the structural state of DNA can be strongly modulated by mechanical constraints.

It is interesting to note that in contrast to B-to-L and B-to-P DNA transitions, which take place even at low DNA tensions, DNA switching from the B- to S-DNA form happens only when the force, f , exerted on the DNA has a sufficiently large magnitude (≥ 50 pN). Such a distinct behavior of L- and P-DNAs, on the one hand, and S-DNA, on the other, mainly originates from the large difference in the structure of the DNA duplex in these DNA forms. Due to rather small helical repeats ($h_L = 16$ bp and $h_P = 3$ bp), L- and P-DNA have much higher relaxed linking number densities ($\Delta l_{k_0,L} = -0.159$ and $\Delta l_{k_0,P} = 0.237$) compared to that of S-DNA ($h_S = 35$ bp and $\Delta l_{k_0,S} = -0.068$); see Table 1. This leads to strong stabilization of L- and P-DNA at large negative and positive torques, respectively, resulting in their dominance over the S-DNA form. In contrast, elongated S-DNA, whose contour length is ~ 1.7 -times larger than that of B-DNA, is more strongly stabilized by DNA stretching, explaining the occurrence of this DNA structure and the L-to-S transition at high forces.

Figure 4c,f showing torque-extension and torque-superhelical density curves reveal further details regarding the DNA structural transitions between different forms. 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 the L-DNA state at sufficiently large negative torques. A further increase of the applied force beyond $f_0 \sim 15$ pN results in the appearance of the B-to-P transition, which takes place at ~ 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 5). 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 an $\sim 50\%$ 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, P or S), they were defined as the set of points (f_0, τ_0) at which $\sim 50\%$ of the DNA segments are in L-, P-, or S-DNA states, respectively. For this purpose, the total number of DNA segments, N_u , in state $u = L, P$ or S at various values of the applied force ($f = f_0$) and torque ($\tau = \tau_0$) was calculated as $(N_u)_{f_0, \tau_0} = -\frac{1}{q} \frac{\partial \ln(Z_{f, \tau})}{\partial \mu_u} \Big|_{f=f_0, \tau=\tau_0}$.

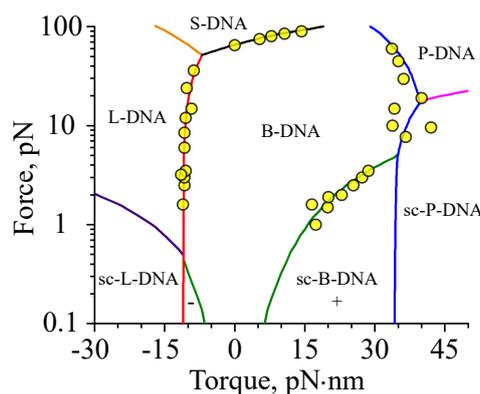


Figure 5. DNA phase diagram. Solid curves indicate transition boundaries between extended (B, L, P, S) and supercoiled (sc-B, sc-L, sc-P) states of DNA predicted by the transfer-matrix calculations. The resulting phase diagram demonstrates good agreement with the existing experimental data (circles), which were digitized from [4–6,8,10,11,22,23].

As can be seen from Figure 5, the resulting phase diagram demonstrates good agreement with existing experimental data, suggesting that the transfer-matrix calculations correctly describe the experimentally observed behavior of DNA [4–6,8,10,11,22,23].

8. Other Potential Applications of the Transfer-Matrix Formalism

In this article, we reviewed several theoretical methods that can be used to investigate the DNA behavior under force and torque constraints, with a special focus on the powerful transfer-matrix approach whose application to the study of the DNA structural transitions has been demonstrated.

It should be noted that while we discussed only the DNA response to torsional stress, it is equally important to understand how DNA behaves when the DNA linking number is subjected to mechanical constraints instead of torque. This type of information may provide additional important insights into the physical processes underlying the DNA structural transitions, as has been demonstrated in a number of previously-reported theoretical studies [51–53,76,84–87]. By using a Dirac delta function approach similar to that described in [52,53,80], it is not hard to extend the transfer-matrix formalism to the fixed DNA linking number scenario, covering a wide range of single-DNA manipulation studies.

Besides the DNA structural transitions, the transfer-matrix calculations can also be employed to investigate DNA-protein interactions under various mechanical constraints. This can be done by treating the protein-bound DNA segments in the discretized polymer model as a new structural state, which is characterized by the protein binding energy, as well as the relaxed linking number change, and bending and twisting rigidities corresponding to the formed nucleoprotein complexes (see the details in [68,88,89] for the case of a mechanically-stretched torsion-free DNA interacting with small DNA-binding proteins). Such an approach allows one to apply the transfer-matrix formalism to gain valuable insights into physical processes governing DNA-protein interactions.

More importantly, this method can be used to solve the inverse problem, as well; by fitting the data collected in single-DNA manipulation experiments, one can obtain detailed information about the physical properties of nucleoprotein complexes formed on DNA. Indeed, upon binding to DNA, proteins often cause certain types of DNA deformations, such as bending or chiral wrapping, resulting in changes to the bending/twisting rigidities, as well as the relaxed linking number of the protein-bound DNA, which can be measured in experiments [31–42,68,90–92]. By fitting the experimental DNA force-extension and/or torque-extension data points to the theoretical curves predicted by the transfer-matrix calculations, one can determine the elastic properties of the formed nucleoprotein complexes and estimate the apparent equilibrium dissociation constant of the studied protein from DNA as a function of force and torque constraints exerted on the DNA.

Furthermore, the capability of the transfer-matrix calculations to take into account local sequence-dependent inhomogeneities of the DNA properties makes it possible to use this method to predict the average occupancies of various DNA sites by a protein that preferentially binds to certain sequence motifs in the context of DNA subjected to various mechanical constraints. Such an approach can be further extended to study competitive binding of multiple protein species to the same DNA and its potential modulation by DNA twisting and stretching. This type of theoretical research will be important in the near future as it may help to gain deeper understanding of the molecular mechanisms involved in the gene regulation, especially in light of recent experimental observations showing that mechanical forces applied to the nucleus can be sensed and processed by living cells, which can interpret the received mechanical signal, activating a number of specific genes in response to it [44–47].

In summary, the flexibility and advantages of the transfer-matrix formalism discussed above make it a powerful tool for a broad range of future applications, including, but not limited, to investigation of sequence-dependent DNA mechanical response, the study of local DNA structural transitions and the exploration of interactions between DNA and architectural proteins.

Acknowledgments: 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 J.Y.

Author Contributions: Artem K. Efremov, Ricksen S. Winardhi and Jie Yan wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

DNA	deoxyribonucleic acid
bp	base-pair
FJC model	freely-jointed chain model
WLC model	worm-like chain model
sc-B, sc-L and sc-P	supercoiled B-, L- and P-DNA, respectively
B-to-L, B-to-P, B-to-S and L-to-S	structural transitions between the corresponding DNA states (B, L, P and S)

References

- Franklin, R.E.; Gosling, R.G. Molecular configuration in sodium thymonucleate. *Nature* **1953**, *171*, 740–741.
- Wang, J.C. Helical repeat of DNA in solution. *Proc. Natl. Acad. Sci. USA* **1979**, *76*, 200–203.
- Rich, A.; Nordheim, A.; Wang, A.H.-J. The chemistry and biology of left-handed Z-DNA. *Ann. Rev. Biochem.* **1984**, *53*, 791–846.
- Bryant, Z.; Stone, M.D.; Gore, J.; Smith, S.B.; Cozzarelli, N.R.; Bustamante, C. Structural transitions and elasticity from torque measurements on DNA. *Nature* **2003**, *424*, 338–341.
- Sheinin, M.Y.; Forth, S.; Marko, J.F.; Wang, M.D. Underwound DNA under tension: Structure, elasticity, and sequence-dependent behaviors. *Phys. Rev. Lett.* **2011**, *107*, 108102.
- Oberstrass, F.C.; Fernandes, L.E.; Bryant, Z. Torque measurements reveal sequence-specific cooperative transitions in supercoiled DNA. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 6106–6111.
- Cluzel, P.; Lebrun, A.; Heller, C.; Lavery, R.; Viovy, J.L.; Chatenay, D.; François, C. DNA: An extensible molecule. *Science* **1996**, *271*, 792–794.
- Smith, S.B.; Cui, Y.; Bustamante, C. Overstretching B-DNA: The elastic response of individual double-stranded and single-stranded DNA molecules. *Science* **1996**, *271*, 795–798.
- Léger, J.F.; Romano, G.; Sarkar, A.; Robert, J.; Bourdieu, L.; Chatenay, D.; Marko, J.F. Structural transitions of a twisted and stretched DNA molecule. *Phys. Rev. Lett.* **1999**, *83*, 1066–1069.
- Allemand, J.F.; Bensimon, D.; Lavery, R.; Croquette, V. Stretched and overwound DNA forms a Pauling-like structure with exposed bases. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 14152–14157.
- Sheinin, M.Y.; Wang, M.D. Twist-stretch coupling and phase transition during DNA supercoiling. *Phys. Chem. Chem. Phys.* **2009**, *11*, 4800–4803.
- Marko, J.F.; Neukirch, S. Global force-torque phase diagram for the DNA double helix: Structural transitions, triple points and collapsed plectonemes. *Phys. Rev. E* **2013**, *88*, 062722.
- Kłysik, J.; Stirdivant, S.M.; Larson, J.E.; Hart, P.A.; Wells, R.D. Left-handed DNA in restriction fragments and a recombinant plasmid. *Nature* **1981**, *290*, 672–677.
- Nordheim, A.; Lafer, E.M.; Peck, L.J.; Wang, J.C.; Stellar, D.; Rich, A. Negatively supercoiled plasmids contain left-handed Z-DNA segments as detected by specific antibody binding. *Cell* **1982**, *31*, 309–318.
- Peck, L.J.; Wang, J.C. Energetics of B-to-Z transition in DNA. *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 6206–6210.
- Strick, T.R.; Croquette, V.; Bensimon, D. Homologous pairing in stretched supercoiled DNA. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 10579–10583.
- Yakovchuk, P.; Protozanova, E.; Frank-Kamenetskii, M.D. Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic Acids Res.* **2006**, *34*, 564–574.
- Fu, H.; Chen, H.; Marko, J.F.; Yan, J. Two distinct overstretched DNA states. *Nucleic Acids Res.* **2006**, *38*, 5594–5600.
- Zhang, X.; Qu, Y.; Chen, H.; Rouzina, I.; Zhang, S.; Doyle, P.S.; Yan, J. Interconversion between three overstretched DNA structures. *J. Am. Chem. Soc.* **2014**, *136*, 16073–16080.
- Franklin, R.E.; Gosling, R.G. The structure of sodium thymonucleate fibres. I. The influence of water content. *Acta Cryst.* **1953**, *6*, 673–677.
- Strick, T.R.; Allemand, J.-F.; Bensimon, D.; Bensimon, A.; Croquette, V. The elasticity of a single supercoiled DNA molecule. *Science* **1996**, *271*, 1835–1837.
- Deufel, C.; Forth, S.; Simmons, C.R.; Dejgosha, S.; Wang, M.D. Nanofabricated quartz cylinders for angular trapping: DNA supercoiling torque detection. *Nat. Methods* **2007**, *4*, 223–225.

23. Forth, S.; Deufel, C.; Sheinin, M.Y.; Daniels, B.; Sethna, J.P.; Wang, M.D. Abrupt buckling transition observed during the plectoneme formation of individual DNA molecules. *Phys. Rev. Lett.* **2008**, *100*, 148301.
24. Fu, H.; Chen, H.; Zhang, X.; Qu, Y.; Marko, J.F.; Yan, J. Transition dynamics and selection of the distinct S-DNA and strand unpeeling modes of double helix overstretching. *Nucleic Acids Res.* **2011**, *39*, 3473–3481.
25. Zhang, X.; Chen, H.; Fu, H.; Doyle, P.S.; Yan, J. Two distinct overstretched DNA structures revealed by single-molecule thermodynamics measurements. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 8103–8108.
26. King, G.A.; Gross, P.; Bockelmann, U.; Modesti, M.; Wuite, G.J.L.; Peterman, E.J.G. Revealing the competition between peeled ssDNA, melting bubbles, and S-DNA during DNA overstretching using fluorescence microscopy. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 3859–3864.
27. Zhang, X.; Chen, H.; Le, S.; Rouzina, I.; Doyle, P.S.; Yan, J. Revealing the competition between peeled ssDNA, melting bubbles, and S-DNA during DNA overstretching by single-molecule calorimetry. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 3865–3870.
28. Mosconi, F.; Allemand, J.F.; Bensimon, D.; Croquette, V. Measurement of the torque on a single stretched and twisted DNA using magnetic tweezers. *Phys. Rev. Lett.* **2009**, *102*, 078301.
29. Vologodskii, A.V.; Cozzarelli, N.R. Conformational and thermodynamic properties of supercoiled DNA. *Annu. Rev. Biophys. Biomol. Struct.* **1994**, *23*, 609–643.
30. Marko, J.F. Biophysics of protein-DNA interactions and chromosome organization. *Physica A* **2015**, *418*, 126–153.
31. Luger, K.; Mäder, A.W.; Richmond, R.K.; Sargent, D.F.; Richmond, T.J. Crystal structure of the nucleosome core particle at 2.8Å resolution. *Nature* **1997**, *389*, 251–260.
32. Skoko, D.; Wong, B.; Johnson, R.C.; Marko, J.F. Micromechanical analysis of the binding of DNA-bending proteins HMGB1, NHP6A, and HU reveals their ability to form highly stable DNA-protein complexes. *Biochemistry* **2004**, *43*, 13867–13874.
33. Van Noort, J.; Verbrugge, S.; Goosen, N.; Dekker, C.; Dame, R.T. Dual architectural roles of HU: Formation of flexible hinges and rigid filaments. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 6969–6974.
34. Swinger, K.K.; Rice, P.A. IHF and HU: Flexible architects of bent DNA. *Curr. Opin. Struct. Biol.* **2004**, *14*, 28–35.
35. Liu, Y.; Chen, H.; Kenney, L.J.; Yan, J. A divalent switch drives H-NS/DNA-binding conformations between stiffening and bridging modes. *Genes Dev.* **2010**, *24*, 339–344.
36. Lin, J.; Chen, H.; Dröge, P.; Yan, J. Physical organization of DNA by multiple non-specific DNA-binding modes of integration host factor (IHF). *PLoS ONE* **2012**, *7*, e49885.
37. Lim, C.J.; Whang, Y.R.; Kenney, L.J.; Yan, J. Gene silencing H-NS paralogue StpA forms a rigid protein filament along DNA that blocks DNA accessibility. *Nucleic Acids Res.* **2012**, *40*, 3316–3328.
38. Winardhi, R.S.; Fu, W.; Castang, S.; Li, Y.; Dove, S.L.; Yan, J. Higher order oligomerization is required for H-NS family member MvaT to form gene-silencing nucleoprotein filament. *Nucleic Acids Res.* **2012**, *40*, 8942–8952.
39. Le, S.; Chen, H.; Cong, P.; Lin, J.; Dröge, P.; Yan, J. Mechanosensing of DNA bending in a single specific protein-DNA complex. *Sci. Rep.* **2013**, *3*, 3508.
40. Qu, Y.; Lim, C.J.; Whang, Y.R.; Liu, J.; Yan, J. Mechanism of DNA organization by *Mycobacterium tuberculosis* protein Lsr2. *Nucleic Acids Res.* **2013**, *41*, 5263–5272.
41. Winardhi, R.S.; Gulvady, R.; Mellies, J.L.; Yan, J. Locus of enterocyte effacement-encoded regulator (Ler) of pathogenic *Escherichia coli* competes off histone-like nucleoid-structuring protein (H-NS) through noncooperative DNA binding. *J. Biol. Chem.* **2014**, *289*, 13739–13750.
42. Efremov, A.K.; Qu, Y.; Maruyama, H.; Lim, C.J.; Takeyasu, K.; Yan, J. Transcriptional repressor TrmBL2 from *Thermococcus kodakarensis* forms filamentous nucleoprotein structures and competes with histones for DNA binding in a salt- and DNA supercoiling-dependent manner. *J. Biol. Chem.* **2015**, *290*, 15770–15784.
43. Ahmad, M.U.D.; Waeger, I.; Hausner, W.; Thomm, M.; Boos, W.; Diederichs, K.; Welte, W. Structural insights into nonspecific binding of DNA by TrmBL2, an archaeal chromatin protein. *J. Mol. Biol.* **2015**, *427*, 3216–3229.
44. Shivashankar, G.V. Mechanosignaling to the cell nucleus and gene regulation. *Annu. Rev. Biophys.* **2011**, *40*, 361–378.
45. Iyer, K.V.; Pulford, S.; Mogilner, A.; Shivashankar, G.V. Mechanical activation of cells induces chromatin remodeling preceding MKL nuclear transport. *Bipolys. J.* **2012**, *103*, 1416–1428.

46. Mammoto, A.; Mammoto, T.; Ingber, D.E. Mechanosensitive mechanisms in transcriptional regulation. *J. Cell Sci.* **2012**, *125*, 3061–3073.
47. Tajik, A.; Zhang, Y.; Wei, F.; Sun, J.; Jia, Q.; Zhou, W.; Singh, R.; Khanna, N.; Belmont, A.S.; Wang, N. Transcription upregulation via force-induced direct stretching of chromatin. *Nat. Mater.* **2016**, *15*, 1287–1296.
48. Marko, J.F.; Siggia, E.D. Stretching DNA. *Macromolecules* **1995**, *28*, 8759–8770.
49. Odijk, T. Stiff chains and filaments under tension. *Macromolecules* **1995**, *28*, 7016–7018.
50. Fain, B.; Rudnick, J.; Östlund, S. Conformations of linear DNA. *Phys. Rev. E* **1997**, *55*, 7364–7368.
51. Moroz, J.D.; Nelson, P. Entropic elasticity of twist-storing polymers. *Macromolecules* **1998**, *31*, 6333–6347.
52. Bouchiat, C.; Mézard, M. Elasticity model of a supercoiled DNA molecule. *Phys. Rev. Lett.* **1998**, *80*, 1556–1559.
53. Bouchiat, C.; Mézard, M. Elastic rod model of a supercoiled DNA molecule. *Eur. Phys. J. E* **2000**, *2*, 377–402.
54. Wang, M.D.; Yin, H.; Landick, R.; Gelles, J.; Block, S.M. Stretching DNA with optical tweezers. *Biophys. J.* **1997**, *72*, 1335–1346.
55. Gross, P.; Laurens, N.; Oddershede, L.B.; Bockelmann, U.; Peterman, E.J.G.; Wuite, G.J.L. Quantifying how DNA stretches, melts and changes twist under tension. *Nat. Phys.* **2011**, *7*, 731–736.
56. Bustamante, C.; Marko, J.F.; Siggia, E.D.; Smith, S. Entropic elasticity of lambda-phage DNA. *Science* **1994**, *265*, 1599–1600.
57. Olson, W.K. Simulating DNA at low resolution. *Curr. Opin. Struct. Biol.* **1996**, *6*, 242–256.
58. Olson, W.K.; Gorin, A.A.; Lu, X.J.; Hock, L.M.; Zhurkin, V.B. DNA sequence-dependent deformability deduced from protein–DNA crystal complexes. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 11163–11168.
59. Lankaš, F.; Šponer, J.; Langowski, J.; Cheatham, T.E., III. DNA basepair step deformability inferred from molecular dynamics simulations. *Biophys. J.* **2003**, *85*, 2872–2883.
60. Lankaš, F.; Gonzalez, O.; Heffler, L.M.; Stoll, G.; Moakher, M.; Maddocks, J.H. On the parameterization of rigid base and basepair models of DNA from molecular dynamics simulations. *Phys. Chem. Chem. Phys.* **2009**, *11*, 10565–10588.
61. Gonzalez, O.; Petkevičiūtė, D.; Maddocks, J.H. A sequence-dependent rigid-base model of DNA. *J. Chem. Phys.* **2013**, *138*, 055102.
62. Călugăreanu, G. Sur les classes d'isotopie des noeuds tridimensionnels et leurs invariants. *Czechoslov. Math. J.* **1961**, *11*, 588–625.
63. White, J.H. Self-linking and the Gauss integral in higher dimensions. *Am. J. Math.* **1969**, *91*, 693–728.
64. Fuller, F.B. The writhing number of a space curve. *Proc. Natl. Acad. Sci. USA* **1971**, *68*, 815–819.
65. Crick, F.H.C. Linking numbers and nucleosomes. *Proc. Natl. Acad. Sci. USA* **1976**, *73*, 2639–2643.
66. Marko, J.F.; Siggia, E.D. Statistical mechanics of supercoiled DNA. *Phys. Rev. E* **1995**, *52*, 2912–2938.
67. Fuller, F.B. Decomposition of the linking number of a closed ribbon: A problem from molecular biology. *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 3557–3561.
68. Yan, J.; Marko, J.F. Effects of DNA-distorting proteins on DNA elastic response. *Phys. Rev. E* **2003**, *68*, 011905.
69. Efremov, A.K.; Winardhi, R.S.; Yan, J. Transfer-matrix calculations of DNA polymer micromechanics under tension and torque constraints. *Phys. Rev. E* **2016**, *94*, 032404.
70. Levitt, M. Protein folding by restrained energy minimization and molecular dynamics. *J. Mol. Biol.* **1983**, *170*, 723–764.
71. Klenin, K.; Langowski, J. Computation of writhe in modeling of supercoiled DNA. *Biopolymers* **2000**, *54*, 307–317.
72. Metropolis, N.; Rosenbluth, A.W.; Rosenbluth, M.N.; Teller, A.H.; Teller, E. Equation of state calculations by fast computing machines. *J. Chem. Phys.* **1953**, *21*, 1087–1092.
73. Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*; Academic Press: San Diego, CA, USA, 2002.
74. Vologodskii, A.V.; Marko, J.F. Extension of torsionally stressed DNA by external force. *Biophys. J.* **1997**, *73*, 123–132.
75. Vologodskii, A.V.; Levene, S.D.; Klenin, K.V.; Frank-Kamenetskii, M.; Cozzarelli, N.R. Conformational and thermodynamic properties of supercoiled DNA. *J. Mol. Biol.* **1992**, *227*, 1224–1243.
76. Marko, J.F. Torque and dynamics of linking number relaxation in stretched supercoiled DNA. *Phys. Rev. E* **2007**, *76*, 021926.

77. Thompson, J.M.T.; van der Heijden, G.H.M.; Neukirch, S. Supercoiling of DNA plasmids: Mechanics of the generalized ply. *Proc. R. Soc. Lond. A* **2002**, *458*, 959–985.
78. Sarkar, A.; Léger, J.-F.; Chatenay, D.; Marko, J.F. Structural transitions in DNA driven by external force and torque. *Phys. Rev. E* **2001**, *63*, 051903.
79. Yan, J.; Marko, J.F. Localized single-stranded bubble mechanism for cyclization of short double helix DNA. *Phys. Rev. Lett.* **2004**, *93*, 108108.
80. Yan, J.; Kawamura, R.; Marko, J.F. Statistics of loop formation along double helix DNAs. *Phys. Rev. E* **2005**, *71*, 061905.
81. Ranjith, P.; Kumar, P.B.S.; Menon, G.I. Distribution functions, loop formation probabilities, and force-extension relations in a model for short double-stranded DNA molecules. *Phys. Rev. Lett.* **2005**, *94*, 138102.
82. Gel'fand, I.M.; Minlos, R.A.; Shapiro, Z.Y. *Representations of the Rotation and Lorentz Groups and Their Applications*; Pergamon Press: New York, NY, USA, 1963.
83. Storm, C.; Nelson, P.C. Theory of high-force DNA stretching and overstretching. *Phys. Rev. E* **2003**, *67*, 051906.
84. Neukirch, S. Extracting DNA twist rigidity from experimental supercoiling data. *Phys. Rev. Lett.* **2004**, *93*, 198107.
85. Purohit, P.K. Plectoneme formation in twisted fluctuating rods. *J. Mech. Phys. Solids* **2008**, *56*, 1715–1729.
86. Daniels, B.C.; Forth, S.; Sheinin, M.Y.; Wang, M.D.; Sethna, J.P. Discontinuities at the DNA supercoiling transition. *Phys. Rev. E* **2009**, *80*, 040901.
87. Marko, J.F.; Neukirch, S. Competition between curls and plectonemes near the buckling transition of stretched supercoiled DNA. *Phys. Rev. E* **2012**, *85*, 011908.
88. Zhang, H.; Marko, J.F. Intrinsic and force-generated cooperativity in a theory of DNA-bending proteins. *Phys. Rev. E* **2010**, *82*, 051906.
89. Zhang, H.; Marko, J.F. Range of interaction between DNA-bending proteins is controlled by the second-longest correlation length for bending fluctuations. *Phys. Rev. Lett.* **2010**, *109*, 248301.
90. Fu, H.; Le, S.; Muniyappa, K.; Yan, J. Dynamics and regulation of RecA polymerization and de-polymerization on double-stranded DNA. *PLoS ONE* **2013**, *8*, e66712.
91. Lee, S.Y.; Lim, C.J.; Dröge, P.; Yan, J. Regulation of bacterial DNA packaging in early stationary phase by competitive DNA binding of Dps and IHF. *Sci. Rep.* **2013**, *5*, 18146.
92. Winardhi, R.S.; Yan, J.; Kenney, L.J. H-NS regulates gene expression and compacts the nucleoid: Insights from single-molecule experiments. *Biophys. J.* **2015**, *109*, 1321–1329.



© 2017 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).