DNA elasticity from coarse-grained simulations: the effect of groove asymmetry

Enrico Skoruppa,† Michiel Laleman,† Stefanos K. Nomidis,†‡ and Enrico Carlon*,†

†KU Leuven, Institute for Theoretical Physics, Celestijnenlaan 200D, 3001 Leuven, Belgium
‡Flemish Institute for Technological Research (VITO), Boeretang 200, B-2400 Mol, Belgium

E-mail: enrico.carlon@kuleuven.be

Abstract

It is well-established that many physical properties of DNA at sufficiently long length scales can be understood by means of simple polymer models. One of the most widely used elasticity models for DNA is the twistable worm-like chain (TWLC), which describes the double helix as a continuous elastic rod with bending and torsional stiffness. An extension of the TWLC, which has recently received some attention, is the model by Marko and Siggia, who introduced an additional twist-bend coupling, expected to arise from the groove asymmetry. By performing computer simulations of two available versions of oxDNA, a coarse-grained model of nucleic acids, we investigate the microscopic origin of twist-bend coupling. We show that this interaction is negligible in the oxDNA version with symmetric grooves, while it appears in the oxDNA version with asymmetric grooves. Our analysis is based on the calculation of the covariance matrix of equilibrium deformations, from which the stiffness parameters are obtained. The estimated twist-bend coupling coefficient from oxDNA simulations is \( G = 30 \pm 1 \) nm, in agreement with the value obtained from a recent analysis of magnetic tweezers data \( (G = 40 \pm 10 \) nm).

Introduction

Owing to its role as the carrier of genetic information, DNA is of central importance in biology. In its interactions with other biomolecules within the cell, DNA is often bent and twisted. A good mechanical model of DNA is therefore essential to understand the complex biological processes in which it is involved. A large number of experiments in the past have shown that its mechanical response can be described using simple continuous polymer models, such as the twistable worm-like chain (TWLC), which treats DNA as an elastic rod, exhibiting resistance to applied bending and twisting. In spite of its simplicity, the TWLC has proven to be surprisingly accurate in the description of the DNA response to applied forces and torques.

As experimental techniques become more accurate, physical models are put to increasingly strict tests. Single-molecule experiments of the past few years have reported some discrepancies between the TWLC predictions and the observed torsional response of DNA. These experiments use magnetic tweezers in order to apply both a torque and a stretching force to a single DNA molecule. The measured torsional stiffness as a function of the applied force turned out to deviate from the TWLC predictions. A recent study explained these discrepancies using an elastic DNA model, which ex-
tends the TWLC by including a direct coupling term between the twisting and bending degrees of freedom. The existence of twist-bend coupling was predicted by Marko and Siggia. Quite surprisingly the consequence of this coupling on the structural and dynamical properties of DNA has only been discussed in a very limited number of papers so far.

In this paper we investigate the elastic properties of oxDNA, a coarse-grained model for simulations of single- and double-stranded DNA. OxDNA comes in two versions: the original version (oxDNA1) contains symmetric grooves, whereas in a more recent extension (oxDNA2) distinct major and minor grooves were introduced. By comparing the two versions, we deduce the effect of an asymmetric grooving on the elastic properties of the molecule. Our analysis shows a clear signature of twist-bend coupling in oxDNA2, while this interaction is absent in the symmetric oxDNA1. This confirms the predictions of Marko and Siggia and shows that the groove asymmetry strongly affects the elastic properties of the molecule. Our estimate of the twist-bend coupling constant in oxDNA2 is in agreement with that obtained from a recent analysis of magnetic tweezers data.

Models and simulations

Elasticity models

Elastic polymer models describe double-stranded DNA as a continuous inextensible rod. At every point along the molecule one defines a local frame of reference, given by a set of three orthonormal vectors \( \{ \hat{e}_1(s), \hat{e}_2(s), \hat{e}_3(s) \} \), where \( 0 \leq s \leq L \) is the arc-length coordinate and \( L \) the contour length. The common convention is to choose \( \hat{e}_3 \) as local tangent to the curve (see Fig. 1), whereas \( \hat{e}_1 \) and \( \hat{e}_2 \) lie in the plane of the ideal, planar Watson-Crick base pairs. The vector \( \hat{e}_1 \) is directed along the symmetry axis of the two grooves and \( \hat{e}_2 \) is obtained from the relation \( \hat{e}_2 = \hat{e}_3 \times \hat{e}_1 \). Knowing how the set \( \{ \hat{e}_1(s), \hat{e}_2(s), \hat{e}_3(s) \} \) depends on \( s \) allows one to reconstruct the conformation of the molecule.

Figure 1: DNA can be represented as an inextensible, twistable, elastic rod. Its conformation is described by a local orthonormal frame, associated with every point along the molecule. \( \hat{e}_3 \) is the unit tangent vector, whereas \( \hat{e}_1 \) is chosen to lie on the symmetry plane of the grooves. The third vector is given by \( \hat{e}_2 = \hat{e}_3 \times \hat{e}_1 \).

Any local deformation of the curve induces a rotation of the frame \( \{ \hat{e}_1, \hat{e}_2, \hat{e}_3 \} \) from \( s \) to \( s + ds \), which can be described by the following differential equation

\[
\frac{d\hat{e}_\mu}{ds} = (\Omega + \omega_0 \hat{e}_3) \times \hat{e}_\mu, \tag{1}
\]

where \( \mu = 1, 2, 3 \) and \( \omega_0 \approx 1.85 \text{ nm}^{-1} \) is the intrinsic twist of the DNA double helix. The vector \( \Omega + \omega_0 \hat{e}_3 \) is parallel to the axis of rotation from \( \hat{e}_\mu(s) \) to \( \hat{e}_\mu(s + ds) \). Note that in general \( \Omega(s) \) depends on the coordinate \( s \). Decomposing this vector along the local frame we define its three components as \( \Omega_\mu(s) \equiv \Omega \cdot \hat{e}_\mu(s) \). The case \( \Omega = |\Omega| \hat{e}_3 \) corresponds to a pure twist deformation, whereas \( \Omega = |\Omega| \hat{e}_1 \) and \( \Omega = |\Omega| \hat{e}_2 \) express bending in the planes defined by \( \hat{e}_1 \) and \( \hat{e}_2 \), respectively.

The lowest-energy configuration of the system is that of zero mechanical stress \( \Omega_1 = \Omega_2 = \Omega_3 = 0 \), which corresponds to a straight rod with an intrinsic twist angle per unit length equal to \( 2\pi\omega_0 \). Expanding around this ground state, one obtains the elastic energy to lowest order in the deformation parameters \( \Omega_\mu \) as

\[
\beta E = \frac{1}{2} \int_0^L \sum_{\mu, \nu = 1}^3 \Omega_\mu(s) M_{\mu\nu} \Omega_\nu(s) ds, \tag{2}
\]

where \( \beta \equiv 1/k_B T \) is the inverse temperature. The \( 3 \times 3 \) symmetric matrix \( M_{\mu\nu} \), which we refer to as the stiffness matrix, contains the elastic constants. Note that from Eq. (1) the \( \Omega \)'s

\[
\hat{e}_1 \hat{e}_2 \hat{e}_3
\]
have the dimension of inverse length. As the left-hand side of Eq. (2) is dimensionless, the elements of the stiffness matrix have the dimension of length. In this work sequence-dependent effects will be neglected, therefore $M$ will not depend on $s$.

Marko and Siggia argued that, due to the asymmetry introduced by the major and minor grooves, the elastic energy of DNA should be invariant only under the transformation $\Omega_1 \rightarrow -\Omega_1$. This implies that $\Omega_2 \Omega_3$ is the only cross-term allowed by symmetry, therefore the stiffness matrix in the Marko-Siggia (MS) model becomes

$$M_{\text{MS}} = \begin{pmatrix} A_1 & 0 & 0 \\ 0 & A_2 & G \\ 0 & G & C \end{pmatrix}, \quad (3)$$

where $A_1 \equiv M_{11}$, $A_2 \equiv M_{22}$, $C \equiv M_{33}$ and $G \equiv M_{23} = M_{32}$. $A_1$ and $A_2$ express the energetic cost of a bending deformation about the local axes $\hat{e}_1$ and $\hat{e}_2$, respectively. $C$ is the intrinsic torsional stiffness, whereas $G$ quantifies the twist-bend coupling interaction. Note that $G \neq 0$ is a direct consequence of the groove asymmetry in the DNA double helix. If one neglects this asymmetry, the MS model reduces to the TWLC model ($G = 0$), and the corresponding stiffness matrix becomes diagonal

$$M_{\text{TWLC}} = \begin{pmatrix} A_1 & 0 & 0 \\ 0 & A_2 & 0 \\ 0 & 0 & C \end{pmatrix}. \quad (4)$$

Most studies model DNA as an isotropic TWLC, for which $A_1 = A_2$.10

**Computer simulations with oxDNA**

In this paper we investigate the elastic properties of oxDNA, which is a model for coarse-grained computer simulations of both single- and double-stranded DNA.13 The model describes DNA as a string of rigid nucleotides, with pairwise interactions modeling the backbone covalent bonds, the hydrogen bonding, the stacking, cross-stacking and excluded-volume interactions. oxDNA has been used in the past for the study of a variety of DNA properties.13–16

We performed simulations using two available versions of the model. The first version (oxDNA1) describes DNA as a molecule with no distinction between major and minor grooves,15 while the second (oxDNA2) introduces distinct grooving asymmetry.14 Figure 2 illustrates molecular conformations of the two models, including a cross-sectional view. As discussed above, the presence of distinct major and minor grooves breaks a molecular symmetry, so we expect that oxDNA1 and oxDNA2 will be mapped onto the TWLC (Eq. (4)) and onto the MS model (Eq. (3)), respectively.

To sample equilibrium fluctuations, molecular dynamics simulations in the NVE ensemble with an Anderson-like thermostat were used. This is implemented in repeated cycles in which the system is first evolved by integrating the Newton equation of motions in time for a given number of steps. Then the momenta of some randomly selected particles are chosen from a Maxwell distribution with a desired simulation temperature ($T = 295$ K in our case). The cycle then repeats itself a large number of times.

Figure 2: Snapshots of configurations of oxDNA1 (top) and oxDNA2 (bottom), including a cross-section view of the helix. While the grooves are symmetric in oxDNA1, distinct major and minor grooves are present in oxDNA2.
Molecular dynamics simulations were performed on 150 basepair molecules using averaged base pair interaction coefficients. In all simulations the salt concentration was set to 0.5M. A total of $10^6$ timesteps were sampled using a numerical integration time step of 15.2 fs and the trajectories were recorded every $5 \times 10^4$ time steps.

**Extraction of elastic parameters**

The pivotal objective of the extraction of elastic parameters is to map oxDNA onto the described elastic model in such a way, that both the elastic properties at the base pair level as well as long range behavior, such as bending and torsional persistence lengths, are captured as accurately as possible. Establishing an appropriate one-to-one correspondence requires the reduction of both models to the same level of complexity. For the continuous elastic model this implies the discretization of the elastic free energy function Eq. (2) to the base-pair level

$$\beta E = \frac{a}{2} \sum_{n=1}^{N} \left( \sum_{\mu, \nu = 1}^{3} \Omega_{\mu}(n) M_{\mu \nu} \Omega_{\nu}(n) \right), \quad (5)$$

where $a = 0.34$ nm is the mean distance between successive base pairs. In the discrete version the deformations $\Omega_{\mu}$ quantify (see Eq. (1)) the finite rotation of a local frame of reference (triad) $\{ \hat{e}_1(n), \hat{e}_2(n), \hat{e}_3(n) \}$, associated with the spatial orientation of the $n$-th base pair of the molecule, into the sequentially adjacent triad $\{ \hat{e}_1(n+1), \hat{e}_2(n+1), \hat{e}_3(n+1) \}$. Therefore an appropriate triad has to be assigned to each base-pair of the oxDNA model, which contains a certain degree of ambiguity, resulting in different mappings for different triads. Such an ambiguity regarding the definition of the tangent vector $\hat{e}_3$ in coarse-grained simulations of DNA and the related implications for the extraction of the bending persistence length have for instance been discussed by Fathizadeh et al., who showed that, when considering short length scales, different definitions of the local tangent vector will usually yield significantly different results for the bending persistence length. However, when considering longer length scales, i.e. comparing more distant tangent vectors, those discrepancies vanish asymptotically.

For a detailed discussion of different triad definitions we refer to the Supporting Information. All results presented in the main text are calculated with a triad definition employing local tangents $\hat{e}_3$ obtained from the mean vector of the intrinsic orientation of the two nucleotides in each basepair, provided by the oxDNA output. The unit vector $\hat{e}_2$ is obtained from the projection of the connecting vector between the centers of the two nucleotides $\mathbf{y}$, onto the orthogonal space of $\hat{e}_3$. Having identified $\hat{e}_3$ and $\hat{e}_2$ we defined $\hat{e}_1 = \hat{e}_2 \times \hat{e}_3$.

In order to infer the stiffness matrix from simulations, we used the standard procedure (see e.g. Ref. 11) which relies on the equipartition theorem\cite{18}

$$\left\langle \Omega_{\mu}(n) \frac{\partial \beta E}{\partial \Omega_{\nu}(n)} \right \rangle = \delta_{\mu \nu}, \quad (6)$$

where $\langle \cdot \rangle$ indicates the thermal average. We introduce next the $3 \times 3$ covariance matrix with elements

$$\Lambda_{\mu \nu} \equiv \left\langle \Omega_{\mu}(n) \Omega_{\nu}(n) \right \rangle, \quad (7)$$

where we have dropped the index $n$ from $\Lambda$, as we neglect sequence-dependent effects. Combining (5) and (6) we get

$$M = \frac{1}{a} \Lambda^{-1}. \quad (8)$$

Thus, the stiffness parameters contained in $M$ can be extracted from the correlation matrix $\Lambda$, obtained from equilibrium fluctuations (Eq. (7)).

This procedure is based on the elastic energy being given by Eq. (5), which in turn assumes that there are no correlations between different sets of $\Omega$’s. To investigate the effect of correlations we introduce the matrix

$$\Xi_{\mu \nu}(m) \equiv \left\langle \left[ \sum_{k=n}^{n+m-1} \Omega_{\mu}(k) \right] \left[ \sum_{l=n}^{n+m-1} \Omega_{\nu}(l) \right] \right \rangle. \quad (9)$$

If correlations beyond neighboring bases are weak, the cross-terms in the previous expres-
sion can be neglected and we obtain

$$
\Xi_{\mu\nu}(m) \approx \sum_{k=n}^{n+m-1} \langle \Omega_{\mu}(k)\Omega_{\nu}(k) \rangle = m \Lambda_{\mu\nu}.
$$

(10)

We define now the $m$-step stiffness matrix as

$$
M(m) \equiv \frac{m}{a} [\Xi(m)]^{-1},
$$

(11)

from which we can obtain the $m$-step elastic constants. In absence of correlations, this matrix will not depend on $m$.

### Results

We present here the results of the simulations highlighting the differences in elastic properties between oxDNA1 and oxDNA2.

#### Probability Distributions

Qualitative evidence of the presence of a non-zero twist-bend coupling in the energy functionals can already be inferred from the distribution of the off-diagonal terms $\Omega_{\mu}(n)\Omega_{\nu}(n)$. Figure 3 shows histogram of these quantities, obtained from simulations of oxDNA1 and oxDNA2. The data are averaged over all base pairs along the DNA contour, hence we drop the position index $n$. While the distribution of $\Omega_1\Omega_2$ and $\Omega_1\Omega_3$ is symmetric and very similar in oxDNA1 and oxDNA2, there is a marked difference between the two models in the histograms of $\Omega_2\Omega_3$. More specifically, in oxDNA1 the distribution appears to be symmetric, whereas in oxDNA2 there is a clear asymmetry, which suggests the existence of a coupling between those deformation parameters.

**Stiffness Matrix** In order to quantify the observed twist-bend coupling interaction, we computed the $m$-step stiffness matrix $M(m)$, as defined in Eq. (11), for both models and for different values of $m$. The results are shown in Fig. 4 where the elements of $M(m)$ are plotted as a function of $m$. In both models the diagonal elements $A_1$, $A_2$ and $C$, as defined in Eqs. (3) and (4), have distinct, non-vanishing values. There is, however, a remarkable difference between oxDNA1 and oxDNA2 in the values of the off-diagonal elements $G$, $M_{12}$ and $M_{13}$. In particular, all off-diagonal elements in oxDNA1 are orders of magnitude smaller when compared to the diagonal ones. On the other hand, although $M_{12}$ and $M_{13}$ remain negligibly small, twist-bend coupling $G$ in oxDNA2 becomes comparable in magnitude to the diagonal terms, which clearly has to be attributed to the
Figure 4: Elastic parameters, obtained from the $m$-step stiffness matrix, as a function of the basepair distance $m$. The remarkable difference between these two sets is the appearance of a significant twist-bend coupling term $G$ for oxDNA2, in contrast to its negligible value in oxDNA1. This is in agreement with the original prediction of Marko and Siggia.\textsuperscript{10}

Asymmetry of the helical grooves. These results are in line with the predictions of Marko and Siggia\textsuperscript{10} and remain valid regardless of the exact choice of coordinate systems (see Supporting Information).

As discussed in the previous section, in absence of correlations between different sets of $\Omega$’s, the elements of $\mathbf{M}(m)$ are expected to be independent of $m$. The results of Fig. 4, however, show that this is not exactly true, which is a signature of the influence of correlations between base pairs separated by more than one nucleotide (though the convergence to a limiting value for increasing $m$ is quite rapid).

When comparing the results among different choices of frames, we find that, despite the different values for $m = 0$, at large $m$ all values are close to each other (see Supporting Information). We, thus, consider these limiting values to be good estimates for the stiffness parameters of the elastic model, onto which oxDNA is mapped. Table 1 summarizes the estimated values of the elastic parameters, averaged over the different choices of local frames, where the error bars reflect the uncertainty from estimates obtained from three different definitions of frames. The first two rows in Table 1 are data obtained from oxDNA simulations in this work, while the last row shows the parametrization obtained from fits of the MS model to magnetic tweezers data.\textsuperscript{9} oxDNA2 data for $C$ and $G$ are consistent with the latter, while some differences are found in $A_1$ and $A_2$. It should be noted, however, that the fitting procedure used in Ref. 9 was not very sensitive to the specific choice of $A_1$ and $A_2$, as other choices fitted the experimental data equally well. The overall quantitative agreement between the oxDNA2 parameters and those from this recent study supports the choice of the plateau values in Fig. 4 as an estimate for the elastic parameters.

The value obtained for $C$ is in general good agreement with previous estimates for oxDNA, which were obtained from methods not involv-
Figure 5: Blue lines: plots of \( l_b \) and \( l_t/2 \) obtained from oxDNA simulations using Eqs. (13) and (14). Red and green lines: analytical predictions for the same quantities in the TWLC (Eqs. (15) and (16)) and in the MS model (Eqs. (17) and (18)), where the \( m \)-dependent stiffnesses of Fig. 4 were used.

|                  | \( A_1 \) | \( A_2 \) | \( C \)   | \( G \)   |
|------------------|----------|----------|----------|----------|
| oxDNA1           | 84(14)   | 29(2)    | 118(1)   | 0.1(0.2) |
| oxDNA2           | 81(10)   | 39(2)    | 105(1)   | 30(1)    |
| Nomidis et al.   | 66(1)    | 46(1)    | 110(5)   | 40(10)   |

Table 1: Values of the stiffness coefficients for oxDNA1 and oxDNA2 obtained in this work (expressed in nm). The last line shows the values obtained from fitting the MS model to magnetic tweezers data.

The equations used to calculate the stiffness coefficients are as follows:

\[
\langle \hat{e}_3(n) \cdot \hat{e}_3(n+m) \rangle \equiv \langle \cos \theta(m) \rangle \sim e^{-ma/\lambda_b}, \tag{12}
\]

where \( \theta(m) \) is the angle formed by the two vectors. As the exponential decay is valid asymptotically in \( m \), we can estimate the bending persistence length from the extrapolation at large \( m \) of the quantity

\[
l_b(m) \equiv -\frac{ma}{\log \langle \cos \theta(m) \rangle}. \tag{13}
\]

Analogously, we can define the twisting persistence length from the decay of the average twist angle

\[
l_t(m) \equiv -\frac{ma}{\log \langle \cos \sum_{k=n}^{n+m} \Omega_3(k) \rangle}. \tag{14}
\]

Equations (13) and (14) can be compared to some analytical expressions. In the TWLC the bending persistence length \( l_b \) is the harmonic
mean of the two bending stiffnesses: \[ l_b = \frac{2A_1A_2}{A_1 + A_2} \] (15)

while the twist persistence length is just twice the torsional stiffness (see e.g. Ref. 23)

\[ l_t = 2C. \] (16)

The same quantities have been calculated for the MS model

\[ l_b = 2A_1 \frac{A_2 - G^2/C}{A_1 + A_2 - G^2/C} \] (17)

and

\[ l_t = 2C \left( 1 - \frac{G^2}{A_2C} \right). \] (18)

From the last two expressions one recovers the TWLC limit upon setting \( G = 0 \).

Figure 5 shows a comparison of the persistence lengths, as obtained from Eq. (13) and (14), with the analytical expressions of the TWLC (Eqs. (15) and (16)) and the MS model (Eqs. (17) and (18)). There is an overall good agreement between the direct computation of the persistence lengths and Eqs. (17) and (18) (with the plateau values of Fig. 4), for both oxDNA1 and oxDNA2. In particular, the prediction of the twisting persistence length is excellent in both models, whereas some small deviations are observed for \( l_b \) (smaller than 10 %). This suggests that there are some features of oxDNA which are not fully captured by the “projection” to an inextensible elastic model, as described by Eq. (2). Note that \( l_b \) in oxDNA2 exhibits a damped oscillatory behaviour at short lengths \( m \) with the helix periodicity, suggesting that the tangent vectors are systematically misaligned. The value of the bending persistence length calculated here is in agreement with previous published oxDNA1 and oxDNA2 data.14,19,20

**Discussion**

Motivated by some recent resurged interest in twist-bend coupling in DNA,9 we have investigated the origin of this interaction in oxDNA, a coarse-grained model of nucleic acids. Twist-bend coupling is a cross-interaction between twist and bending degrees of freedom. In the context of DNA, the existence of such an interaction was predicted by Marko and Siggia,10 who argued that twist-bend coupling follows from the groove asymmetry, a characteristic of the DNA molecular structure. Although this prediction is more than 20 years old, twist-bend coupling has received little attention in the DNA literature. The standard polymer model currently used to describe DNA elasticity remains the TWLC, which is characterized by a diagonal stiffness matrix (Eq. (4)), with no cross-terms.

OxDNA is particularly suited to investigate the origin of twist-bend coupling, as it comes in two different versions (oxDNA1 and oxDNA2). The double helical grooves are symmetric in oxDNA1 and asymmetric in oxDNA2, with widths reproducing the average B-DNA geometry. Our simulations, sampling equilibrium conformations of both oxDNA1 and oxDNA2, show that only the latter model has a significant twist-bend coupling term (Fig. 4). This is in agreement with the symmetry argument by Marko and Siggia.10

The estimated twist-bend coupling coefficient from oxDNA2 is \( G = 30 \pm 1 \) nm, which agrees with the value \( G = 40 \pm 10 \) nm, obtained from fitting magnetic tweezer data.9 An earlier estimate of \( G \approx 25 \) nm was obtained from the analysis of structural correlations of DNA wrapped around histone proteins12 in addition, all atom simulations also support the existence of a twist-bend coupling term,11,21,24 although those studies are restricted to short fragments (\( \approx 20 \) bp).

One of the most remarkable effects of twist-bend coupling in DNA is the appearance of a novel twist length scale\(^9\) (Eq. (18)) with an associated twist stiffness \( \kappa_t = l_t/2 \), which differs from the intrinsic value \( C \). In the MS model a pure twist deformation (\( \Omega_1 = \Omega_2 = 0 \), \( \Omega_3 \neq 0 \)) has an associated stiffness \( C \). However, equilibrium fluctuations involve all \( \Omega \)'s simultaneously, so the twist stiffness, as observed from equilibrium fluctuations, is influenced by \( G \) (Eq. (18)). From oxDNA2 simulations we es-
timate $\kappa_t = l_t/2 \approx 83$ nm (see Fig. 5). This is close to the value $\kappa_t = 75$ nm, recently obtained from fitting the MS model to magnetic tweezers data.\(^9\)

In the past, several experiments were performed to determine the intrinsic twist stiffness, typically producing values in the range $75 \text{ nm} \lesssim C \lesssim 120$ nm. This wide range of values has a natural explanation within the MS model.\(^9\) The above measurements can be subdivided in two groups. In a first group one can set all techniques which sample freely fluctuating DNA for which the MS model predicts $\kappa_t \approx 80$ nm. Typical experiments on unstretched DNA yield indeed torsional stiffness values around 75 nm. In a second group one can place all techniques which measure DNA under applied force. At sufficiently high force, bending fluctuations are suppressed ($\Omega_1$ and $\Omega_2$ are small) therefore one expects to measure pure twist deformations, which are governed by a stiffness $C \approx 110$ nm. This is indeed in agreement with magnetic tweezers data. The conclusion is that the MS model can account for the wide range of the torsional stiffness measurements.

Having shown that the twist-bend coupling is a relevant interaction in DNA, one can ask in which limits and for which quantities the TWLC can still be considered a good approximated DNA model. Our work shows that one can map freely fluctuating DNA onto a TWLC using $C \approx 80$ nm as twist elastic parameter, which incorporates the effect of twist-bend coupling. However some care needs to be taken in the presence of a stretching force, as the suppression of bending fluctuation will influence the twist stiffness. At high forces DNA will then be mapped onto an effective TWLC with a higher value of $C$. Finally, it will be important to investigate the effect of twist-bend coupling in cases where DNA behavior is influenced by its mechanics as in DNA supercoiling\(^25,26\) or in DNA-protein interactions.\(^27,28\)

**Acknowledgement** Discussions with F. Kriegel, F. Lankäš, J. Lipfert, C. Matek and W. Vanderlinden are gratefully acknowledged. We thank T. Dršata for analyzing the all-atom simulation trajectories\(^21\) from which stiffness data in Supporting Information, Table 3, were obtained. We acknowledge financial support from KU Leuven grant IDO/12/08, and from the Research Funds Flanders (FWO Vlaanderen) grant VITO-FWO.

**References**

1. Bryant, Z.; Oberstrass, F. C.; Basu, A. *Curr. Opin. Struct. Biol.* 2012, 22, 304–312.

2. Nelson, P.; Radosavljevic, M.; Bromberg, S. Biological physics: energy, information, life; W.H. Freeman and Co., New York, 2002.

3. Marko, J. F.; Siggia, E. D. *Macromolecules* 1995, 28, 8759–8770.

4. Bustamante, C.; Bryant, Z.; Smith, S. B. *Nature* 2003; 421, 423–427.

5. Marko, J. F.; Siggia, E. D. *Science* 1994; 265, 506–508.

6. Strick, T.; Allemand, J.-F.; Bensimon, D.; Bensimon, A.; Croquette, V. *Science* 1996; 271, 1835–1837.

7. Lipfert, J.; Kerssemakers, J. W.; Jager, T.; Dekker, N. H. *Nat. Methods* 2010, 7, 977–980.

8. Lipfert, J.; Wiggin, M.; Kerssemakers, J. W.; Pedaci, F.; Dekker, N. H. *Nat. Commun.* 2011, 2, 439.

9. Nomidis, S. K.; Kriegel, F.; Vanderlinden, W.; Lipfert, J.; Carlon, E. Manuscript submitted for publication.

10. Marko, J.; Siggia, E. *Macromolecules* 1994, 27, 981–988.

11. Lankäš, F.; Šponer, J.; Hobza, P.; Langowski, J. *J. M. Biol.* 2000, 299, 695–709.

12. Mohammad-Rafiee, F.; Golestanian, R. *Phys. Rev. Lett.* 2005, 94, 238102.
(13) Ouldridge, T. E.; Louis, A. A.; Doye, J. P. 
*Phys. Rev. Lett.* **2010**, *104*, 178101.

(14) Snodin, B. E.; Randisi, F.; Mosayebi, M.; Šulc, P.; Schreck, J. S.; Romano, F.; Ouldridge, T. E.; Tsukanov, R.; Nir, E.; Louis, A. A. *J. Chem. Phys.* **2015**, *142*, 234901.

(15) Šulc, P.; Romano, F.; Ouldridge, T. E.; Rovigatti, L.; Doye, J. P. K.; Louis, A. A. *J. Chem. Phys.* **2012**, *137*, 135101.

(16) Sutthibutpong, T.; Matek, C.; Benham, C.; Slade, G. G.; Noy, A.; Laughton, C.; Doye, J. P.; Louis, A. A.; Harris, S. A. *Nucl. Acids Res.* **2016**, *44*, 9121–9130.

(17) Fathizadeh, A.; Eslami-Mossallam, B.; Ejtehadi, M. R. *Phys. Rev. E* **2012**, *86*, 051907.

(18) Huang, K. *Statistical Mechanics*; J. Wiley, 1987.

(19) Ouldridge, T. E.; Louis, A. A.; Doye, J. P. *J. Chem. Phys.* **2011**, *134*, 085101.

(20) Matek, C.; Ouldridge, T. E.; Doye, J. P.; Louis, A. A. *Scientific Reports* **2015**, *5*, 7655.

(21) Lankaš, F.; Šponer, J.; Langowski, J.; Cheatham, T. E. *Biophys J.* **2003**, *85*, 2872–2883.

(22) Eslami-Mossallam, B.; Ejtehadi, M. *Phys. Rev. E* **2009**, *80*, 011919.

(23) Brackley, C.; Morozov, A.; Marenduzzo, D. *J. Chem. Phys.* **2014**, *140*, 135103.

(24) Dršata, T.; Špačková, N.; Jurečka, P.; Zgarbová, M.; Šponer, J.; Lankaš, F. *Nucl. Acids Res.* **2014**, *42*, 7383–7394.

(25) Lepage, T.; Képès, F.; Junier, I. *Biophys. J.* **2015**, *109*, 135–143.

(26) Fathizadeh, A.; Schiessel, H.; Ejtehadi, M. *Macromolecules* **2015**, *48*, 164–172.

(27) Becker, N. B.; Everaers, R. *J. Chem. Phys.* **2009**, *130*, 04B602.

(28) Marko, J. F. *Physica A* **2015**, *418*, 126–153.
Triad Definitions

Continuous Chain

In order to describe any local deformations of an inextensible, elastic rod, onto which DNA can be mapped, one has to introduce a local frame of reference \( \{ \hat{e}_1(s), \hat{e}_2(s), \hat{e}_3(s) \} \) (triad) to every point along the rod. The deformations can be, thus, determined from the rotation of one triad into the next one. In the case of a continuous chain, the following differential equation will hold

\[
\frac{d\hat{e}_\mu}{ds} = (\Omega + \omega_0 \hat{e}_3) \times \hat{e}_\mu
\]

and this frame of reference can be unambiguously defined: \( \hat{e}_3 \) may be taken to be the tangent to the curve, \( \hat{e}_1 \) pointing along the symmetry axis of the two grooves (oriented towards the major groove) and \( \hat{e}_2 \) simply given by \( \hat{e}_2 = \hat{e}_3 \times \hat{e}_1 \).

oxDNA

In the discrete case of oxDNA, different triads can be defined using the few reference points provided by the coarse-grained model. In particular, oxDNA consists of rigid nucleotides represented by three interactions sites: the hydrogen-bonding, stacking and backbone sites (T. Ouldridge, PhD Thesis, University of Oxford (2011)). The orientation of each nucleotide is given by a normal vector \( \hat{n} \), specifying the plane of the base, and a vector \( \hat{b} \) pointing from the stacking site to the hydrogen-bonding site (as in Fig. 1). For oxDNA1 all three sites lie on the same straight line, while in oxDNA2 the position of the backbone site is changed, thus inducing the grooving asymmetry (B.E. Snodin et al. J. Chem. Phys. 142, 234901 (2015)). Hence each base-pair comes with 2 intrinsic triads (one per nucleotide), with the normal vectors pointing in the respective 5'-3' direction of the strands. The interactions are designed such that in the minimum energy configuration the vectors \( \hat{b}_{\text{nuc}1} \) and \( \hat{b}_{\text{nuc}2} \), attached to the two nucleotides of the same base-pair, point directly towards each other.

In what follows we present the four different choices of triads we have tested.

**Triad I.** The aforementioned intrinsic nucleotide triads present a natural definition for the triad attached to a base pair. The base-pair normal vector can be constructed as the average vector of
the nucleotide normal vectors
\[ \hat{e}_3 = \frac{\hat{n}_{\text{nuc}1} - \hat{n}_{\text{nuc}2}}{\|\hat{n}_{\text{nuc}1} - \hat{n}_{\text{nuc}2}\|}. \] (2)

The mean vector of \( \hat{b}_{\text{nuc}1} \) and \( \hat{b}_{\text{nuc}2} \)
\[ \hat{y} = \frac{\hat{b}_{\text{nuc}1} - \hat{b}_{\text{nuc}2}}{\|\hat{b}_{\text{nuc}1} - \hat{b}_{\text{nuc}2}\|} \] (3)

can be approximately identified with \( \hat{e}_2 \), however in general it will fail to be orthogonal to \( \hat{e}_3 \). This can easily be rectified by projecting it onto the orthogonal space of \( \hat{e}_3 \)
\[ \hat{e}_2 = \frac{\hat{y} - (\hat{y} \cdot \hat{e}_3)\hat{e}_3}{\|\hat{y} - (\hat{y} \cdot \hat{e}_3)\hat{e}_3\|}. \] (4)

The last vector is simply given by \( \hat{e}_1 = \hat{e}_2 \times \hat{e}_3 \).

**Triad II.** Alternatively, \( \hat{e}_2 \) can be obtained from connecting the centers of mass \( r_{\text{nuc}1} \) and \( r_{\text{nuc}2} \) of the two nucleotides
\[ \hat{y} = \frac{r_{\text{nuc}1} - r_{\text{nuc}2}}{\|r_{\text{nuc}1} - r_{\text{nuc}2}\|} \] (5)
and find the complete triad in a completely analogous way as for Triad I. This particular choice of triad was used in the main article, as it appeared to be the most robust (i.e. it yielded the smallest correlations between consecutive \( \Omega_{\mu} \)).

**Triad III.** The tangent vector can also be constructed using the center of mass of the nucleotides. The center of mass of the i-th basepair can be defined as
\[ R_{\text{bp}}(i) = \frac{r_{\text{nuc}1}(i) + r_{\text{nuc}2}(i)}{2}. \] (6)

Identifying the normalized connectors of consecutive \( R_{\text{bp}}(i) \) with \( \hat{e}_3 \) would result in a directionally-dependent definition, therefore \( \hat{e}_3 \) was chosen as the connector between the center of masses of the previous and next basepair
\[ \hat{e}_3(i) = \frac{R_{\text{bp}}(i + 1) - R_{\text{bp}}(i - 1)}{\|R_{\text{bp}}(i + 1) - R_{\text{bp}}(i - 1)\|}. \] (7)

The definition of the remaining triad versors is identical to the one used for Triad II.

**Triad IV.** Instead of selecting one vector as the arithmetic mean and projecting the others on its orthogonal space, one can attempt to treat them on a more equal footing. By placing the 3 nucleotide triad vectors in the columns of a matrix one obtains a rotation matrix
\[ T_{\text{nuc}} = [\hat{t}_{\text{nuc}}, \hat{b}_{\text{nuc}}, \hat{n}_{\text{nuc}}] \in SO(3), \] (8)
with \( \hat{t}_{\text{nuc}} = \hat{b}_{\text{nuc}} \times \hat{n}_{\text{nuc}} \). The arithmetic mean \( T = \frac{1}{2} (T_{\text{nuc}1} + T_{\text{nuc}2}) \) will generally not be a rotation matrix itself, it is however possible to orthogonally project \( T \) onto \( SO(3) \). It can be shown that this projection is given by (M. Moakher, SIAM J. Matrix Anal. Appl. 24, 1 (2002))
\[ T = TU \text{ diag} \left( \frac{1}{\sqrt{\Lambda_1}}, \frac{1}{\sqrt{\Lambda_2}}, \frac{1}{s \sqrt{\Lambda_3}} \right) U^\top, \] (9)
where \( \overline{T} = \frac{1}{N} \sum_{k=1}^{N} T^{(k)} \), \( \Lambda_1 \geq \Lambda_2 \geq \Lambda_3 \geq 0 \) are the eigenvalues of \( M = \overline{T}^\top \overline{T} \), the matrix \( U \) defined so that \( U^\top M U = \text{diag}(\Lambda_1, \Lambda_2, \Lambda_3) \). The variable \( s \) is defined so that \( s = 1 \) if \( \det \overline{T} > 0 \) and \( s = -1 \) if \( \det \overline{T} < 0 \).

### Calculation of \( \Omega \)

#### Rodrigues’ Formula

A generalization of Eq. (1) (valid for infinitesimal rotations) can be obtained for a general rotation. According to Rodrigues’ rotation formula, the rotation of a vector \( \mathbf{v} \) about an axis \( \hat{\Omega} \) by an angle \( \Omega \) is given by

\[
\mathbf{v}_{\text{rotated}} = \mathbf{v} \cos \Omega + \left( \hat{\Omega} \times \mathbf{v} \right) \sin \Omega + \hat{\Omega} (\hat{\Omega} \cdot \mathbf{v}) (1 - \cos \Omega). \tag{10}
\]

From each triad one can construct an orthogonal matrix, by placing the triad vectors in the columns of a \( 3 \times 3 \) matrix

\[
\mathbf{T}(i) = [\hat{e}_1(i), \hat{e}_2(i), \hat{e}_3(i)] \in SO(3). \tag{11}
\]

This matrix is exactly the rotation matrix, transforming the canonical frame into the frame of the respective triad. The matrix rotating \( \mathbf{T}(i) \) into \( \mathbf{T}(i+1) \) with respect to coordinate system of the \( i \)-th triad is given by

\[
\mathbf{R} = \mathbf{T}^\top(i) \mathbf{T}(i+1). \tag{12}
\]

It is straightforward to show that in this frame the rotation matrix \( \mathbf{R} \) can by written in terms of the components of the rotation vector \( \Omega = (\Omega_1 \ \Omega_2 \ \Omega_3)^\top \)

\[
\mathbf{R}(\Omega) = \left( \begin{array}{ccc}
\cos \Omega + \left( \frac{\Omega_1}{\Omega_2} \right)^2 (1 - \cos \Omega) & \frac{\Omega_1 \Omega_2}{\Omega_2^2} (1 - \cos \Omega) - \frac{\Omega_1}{\Omega_2} \sin \Omega & \frac{\Omega_1 \Omega_2}{\Omega_2^2} (1 - \cos \Omega) + \frac{\Omega_1}{\Omega_2} \sin \Omega \\
\frac{\Omega_1 \Omega_2}{\Omega_2^2} (1 - \cos \Omega) + \frac{\Omega_1}{\Omega_2} \sin \Omega & \cos \Omega + \left( \frac{\Omega_1}{\Omega_2} \right)^2 (1 - \cos \Omega) & \frac{\Omega_1 \Omega_2}{\Omega_2^2} (1 - \cos \Omega) - \frac{\Omega_1}{\Omega_2} \sin \Omega \\
\frac{\Omega_1 \Omega_2}{\Omega_2^2} (1 - \cos \Omega) - \frac{\Omega_1}{\Omega_2} \sin \Omega & \frac{\Omega_1 \Omega_2}{\Omega_2^2} (1 - \cos \Omega) + \frac{\Omega_1}{\Omega_2} \sin \Omega & \cos \Omega + \left( \frac{\Omega_1}{\Omega_2} \right)^2 (1 - \cos \Omega)
\end{array} \right). \tag{13}
\]

The deformation parameters \( \Omega_\mu \) can now be extracted by equating Eqs. (12) and (13) and solving for \( \Omega_1, \ \Omega_2 \) and \( \Omega_3 \). A simple way to do this is by noticing that

\[
\text{tr}(\mathbf{R}) = 1 + 2 \cos \Omega. \tag{14}
\]

Moreover, one can also verify that the following relation holds

\[
\Omega = \frac{\Omega}{2 \sin \Omega} \begin{pmatrix}
R_{32} - R_{23} \\
R_{13} - R_{31} \\
R_{21} - R_{12}
\end{pmatrix}. \tag{15}
\]

Note that the sign ambiguity presented in Eq. (14) is completely inconsequential for Eq. (15).

### Distributions of \( \Omega \)’s

The approximation of the free energy by a quadratic form

\[
\beta E = \frac{a}{2} \sum_{i=1}^{N} \Omega^\top(i) M \Omega(i). \tag{16}
\]
Figure 2: Probability distribution of $\Omega_\mu$ for oxDNA1 and oxDNA2 in logscale, using Triad II. The distributions coincide very well with Gaussian distributions, validating the use of a quadratic form for the free energy. Analogous distributions are found for the other choices of triads.

implicitly assumes that the deformation parameters follow a Gaussian distribution. Figure 2 shows the distributions of $\Omega_1$, $\Omega_2$ and $\Omega_3$ (blue lines) as obtained from equilibrium simulations. For a clear comparison, the distributions are shown in logarithmic scale. The fitted Gaussian curves (green lines) indicate that the quadratic approximation is excellent for $\Omega_1$ and $\Omega_2$, while some small deviations are observed in the distributions of $\Omega_3$ (noticeable for angles larger than 15 degrees). The distributions of $\Omega_3$ are slightly asymmetric, which is a consequence of the intrinsic twist $\omega_0$ (different response of DNA to under- and over-twisting).

Stiffness parameters for alternative triads definitions

The extracted stiffness parameters for the 4 different triads are summarized in Fig. 3 and Table 2. The plateau values (large $m$) are quite consistent among the different triad definitions, with the exception of $A_1$ and $A_2$ obtained from Triad III. On the other hand, the values obtained for $m = 1$ are significantly more diverse.

Stiffness matrices from all-atom simulations

Table 3 shows the elements of one-step stiffness matrices from the all-atom simulations of Lankaš et al. (Biophys. J. 85, 2872 (2003)). The data are in nm and include all 10 possible sequence combinations of neighboring nucleotides along one strand. Although the analysis of Lankaš et al. included various stretching deformations, here we only show the rotational coordinates, by integrating out the translational degrees of freedom. The last column in Table 3 gives the average over the 10 sequence-dependent terms. One notices that $G$ (twist-roll or twist-bend coupling) is the largest among all off-diagonal terms. All data refer to one-step stiffnesses (i.e. they roughly match the $m = 1$ data of Fig. 3 and are significantly smaller than the expected limiting values, as found in oxDNA calculations.)
Figure 3: Elements of the $m$-step stiffness matrix as a function of the base-pair distance $m$, extracted from 4 different triad definitions. Data based on Triad I, II, III and IV are shown with triangles, circles, diamonds and pentagons, respectively. Note that the spread in the plateau values of $G$ and $C$ is remarkably small, despite the large differences at $m = 1$. For the bending stiffness parameters $A_1$ and $A_2$, Triads I, II and IV practically yield the same plateau values, while Triad III tends to give quite different values. This is probably due to the fundamentally different definition of the tangent vector in Triad III.

Table 2: Values of the stiffness coefficients (expressed in nm) for oxDNA1 and oxDNA2 for different Triad definitions. The values given here correspond to the plateau values of Fig. 3.

|        | oxDNA1 | oxDNA2 |
|--------|--------|--------|
|        | $A_1$  | $A_2$  | $C$  | $G$  | $A_1$  | $A_2$  | $C$  | $G$  |
| Triad I| 76     | 30     | 120  | 0.1  | 76     | 40     | 105  | 29.8 |
| Triad II| 75     | 30     | 118  | 0.2  | 75     | 40     | 104  | 29.6 |
| Triad III| 109    | 25     | 118  | -0.3 | 99     | 35     | 106  | 30.7 |
| Triad IV| 75     | 30     | 118  | 0.1  | 75     | 41     | 104  | 29.6 |
Table 3: Elements of the stiffness matrix (expressed in nm) for different base pairs, obtained from all-atom simulations (courtesy of F. Lankaš and T. Dršata). In order to facilitate the readout, we have included the tilt, roll and twist nomenclature, which corresponds to our definition of $\Omega_1$, $\Omega_2$ and $\Omega_3$, respectively.

|                  | CG  | CA  | TA  | AG  | GG  | AA  | GA  | AT  | AC  | GC  | average |
|------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|---------|
| tilt-tilt ($A_1$)| 47.6| 50.6| 44.5| 67.3| 70.7| 60.9| 69.9| 73.6| 75.0| 70.0| 63.0    |
| roll-roll ($A_2$)| 27.7| 31.4| 24.5| 41.0| 44.4| 42.2| 38.7| 45.1| 46.1| 47.3| 38.8    |
| twist-twist ($C$)| 32.7| 34.0| 57.6| 57.9| 58.9| 49.5| 46.6| 77.7| 65.1| 51.7| 53.2    |
| tilt-roll ($M_{12}$)| 2.8 | 1.3 | 0.1 | -5.3| -1.7| 3.6 | -0.2| 0.4 | 4.0 | -0.5| 0.4     |
| tilt-twist ($M_{13}$)| 4.4 | -1.5| -1.1| -3.9| 0.9 | 6.7 | 0.0 | -0.7| -0.6| -0.7| 0.4     |
| roll-twist ($G$)  | 3.7 | 5.8 | 14.1| 6.7 | 7.4 | 10.5| 15.7| 11.9| 13.4| 13.0| 10.2    |