Length-scale-dependent elasticity in DNA from coarse-grained and all-atom models

Enrico Skoruppa,[1] Aderik Voorspoels,[1] Jocelyne Vreede,[2] and Enrico Carlon[1]
[1]*Laboratory for Soft Matter and Biophysics, KU Leuven, Celestijnenlaan 200D, 3001 Leuven, Belgium*
[2]*Van 't Hoff Institute for Molecular Sciences, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, the Netherlands*

(Received 29 September 2020; accepted 25 March 2021; published 12 April 2021)

We investigate the influence of nonlocal couplings on the torsional and bending elasticities of DNA. Such couplings have been observed in the past by several simulation studies. Here, we use a description of DNA conformations based on the variables tilt, roll, and twist. Our analysis of both coarse-grained (oxDNA) and all-atom models indicates that these share strikingly similar features: there are strong off-site couplings for tilt-tilt and twist-twist, while they are much weaker in the roll-roll case. By developing an analytical framework to estimate bending and torsional persistence lengths in nonlocal DNA models, we show how off-site interactions generate a length-scale-dependent elasticity. Based on the simulation-generated elasticity data, the theory predicts a significant length-scale-dependent effect on torsional fluctuations but only a modest effect on bending fluctuations. These results are in agreement with experiments probing DNA mechanics from single base pair to kilobase pair scales.

DOI: 10.1103/PhysRevE.103.042408

I. INTRODUCTION

Mechanical properties of DNA strongly influence how the double helix performs its various tasks in the cell, where it is often bent and twisted [1]. Computer simulations have been playing an increasingly important role in understanding these properties. Depending on the length scale relevant to the particular issue at hand and the level of detail required, simulations of either atomistic [2–8] or coarse-grained resolution [9–21] can be employed. At sufficiently long length scales, the mechanical response of DNA is well described by continuous elastic models, such as the twistable wormlike chain (TWLC), which neglects sequence-dependent effects [22]. The TWLC uses three local configurational variables: tilt and roll (describing the two possible bending modes) and twist. The free energy associated with deformations is quadratic and local, i.e., distal sites are assumed to be independent. We will refer to these types of model as on-site models. Describing DNA at shorter distances requires a more detailed approach since two effects become relevant: sequence specificity starts to dominate the elastic behavior and the assumption of coupling locality no longer holds. The former issue is well-documented—several studies have shown that DNA elasticity at the base-pair level is strongly dependent on the involved type of nucleotides [3–5]—while the latter issue is the main concern of this paper. Couplings beyond nearest neighbors have been observed in all-atom simulations [2,5,23,24] as well as in coarse-grained models [16]. Lankas *et al.* [2], in one of the earliest all-atom simulation study to characterize DNA elastic constants, observed nonlocal correlations which are more pronounced for some degrees of freedom. These effects were also discussed in the context of the rigid base and rigid base-pair models [23]. Noy and Golestanian [5] performed a systematic study of the effective elastic properties of DNA at different length scales. Eslami-Mossallam and Ejtehadi [24] showed, using Monte Carlo simulations, that long-range couplings enhance the local bendability of DNA when compared to the prediction of purely local models and that the bending persistence length becomes length dependent.

The aim of this paper is to analyze mechanical models of DNA with off-site couplings and to incorporate these effects in a homogeneous TWLC-like model of DNA for which we derive analytical expressions of bending and torsional persistence lengths. We assume quadratic interactions both for on-site and off-site terms and do not discuss extreme bendability at short scales and kinking, which would require a more complex energetic model including beyond-harmonic interactions (for a recent study of kinking, see, e.g., Ref. [25]). For the purpose of this paper, which focuses on the behavior of the bending and torsional persistence lengths, a description of the DNA configurations using tilt, roll and twist is sufficient. Notwithstanding, the methodology developed here is applicable to more sophisticated models with an enlarged phase space of collective variables, such as the rigid base-pair model [23].

Here, we present the results of simulations of a homogeneous coarse-grained DNA model and of an all-atom model for which we average over different sequences. The central quantity in our analysis is the set of momentum space stiffness matrices that captures the linear response of the model at all length scales and presents a convenient way to quantify the effect of off-site interactions. The momentum space formulation is natural in the context of our work as we consider the limit of infinitely long homogeneous molecules—implying translational invariance and independence of modes with different momenta. Moreover, length-scale-dependent bending and torsional persistence lengths can be directly deduced by a momentum space integration. The asymptotic value of the

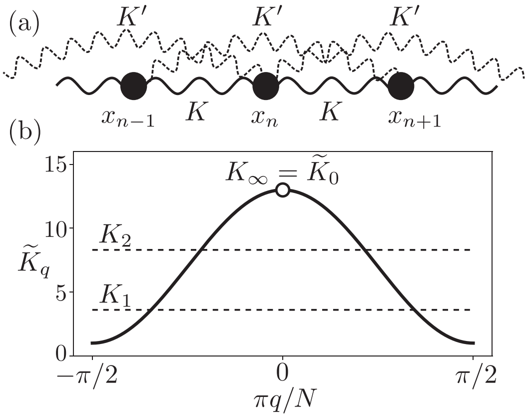

FIG. 1. (a) Toy model of length-scale-dependent elasticity consisting of a linear chain with neighbors and next-neighbor springs with stiffnesses K and K', respectively [Eq. (1)]. (b) Momentum space stiffness of the model (7) for $K = 1$ and $K' = 3$. The one-step K_1, two-step K_2, and asymptotic stiffnesses $K_\infty = \widetilde{K}_0$ [Eqs. (12)–(14)] are shown. In the case shown here ($K' > 0$), the system is softer at short scales: $K_1 < K_2 < \ldots < K_\infty$.

torsional persistence length is obtained from the zero-momentum component, i.e., for $q \to 0$, of the associated stiffness matrices. As we shall show, the intrinsic twist of the DNA chain prompts the asymptotic bending persistence length to be associated with the momentum space components in the limit $q \to \pm\Delta q$, where Δq is the momentum related to one turn of the double helix. Hence, from a plot of stiffnesses in momentum space, one can easily read off the asymptotic values of the persistence lengths and infer the short distance behavior as well. Although our focus here is DNA, it turns out that length-scale-dependent elasticity can also be understood in simpler systems. Therefore, we start our discussion introducing a toy model (Sec. II). This model shows a length-scale-dependent elastic stiffness [Eq. (11)] and the exponential decay of a local perturbation [Eq. (18)] which are also found in DNA. The advantage is that the toy model is simpler and perhaps more intuitive to understand. In addition, several quantities can be computed exactly. In Sec. III, the formalism introduced for the simple model is transferred to our three-dimensional model for DNA. Numerical results obtained with the coarse-grained and atomistic model are presented in Sec. IV. Finally, in Sec. V, we discuss the results obtained and link our findings to experimental observations.

II. LINEAR ELASTIC CHAIN WITH NEXT NEAREST-NEIGHBOR COUPLING

To illustrate the effect of beyond-nearest-neighbor couplings and the procedure of analyzing length-dependent elasticity, we first consider a one-dimensional toy model of a linear elastic chain with next-neighbors couplings. This model [illustrated in Fig. 1(a)] consists of an elastic chain of N masses located at positions x_0, x_1, ...x_{N-1}. Interactions between the masses are mediated by two types of springs with stiffnesses K and K' and rest lengths a and $2a$, acting, respectively, between nearest neighbors and next-nearest neighbors. Accordingly, the energy of the system—in units of k_BT—is given by

$$\beta E = \frac{K}{2}\sum_{n=0}^{N-1}(x_{n+1} - x_n - a)^2 + \frac{K'}{2}\sum_{n=0}^{N-1}(x_{n+2} - x_n - 2a)^2, \tag{1}$$

with $\beta = 1/k_BT$. We use periodic boundary conditions by defining $x_{N+1} \equiv x_N + x_1 - x_0$, which adds an auxiliary spring at one end of the chain having the same amount of stretching as the first spring at the opposite end. These boundary conditions are formally necessary for our formalism, however, their violation merely constitutes a finite-size effect that will vanish for sufficiently large N. The minimal energy configuration of the system is $x_n = x_0 + na$. We are interested in the stretching fluctuations at different length scales, as captured by the m-step fluctuations

$$\langle (x_m - x_0 - ma)^2 \rangle = \frac{m}{K_m}, \tag{2}$$

for which we define an effective spring constant K_m. In absence of next-nearest-neighbor couplings ($K' = 0$), one simply finds $K_m = K$, as the mean-squared extension of m independent springs is just m times the extension of a single spring, which yields the stated relation by virtue of the equipartition theorem. As we shall show, in the case $K' \neq 0$ the spring constant K_m depends on m, indicating a length-dependent elasticity.

For the calculation of K_m, we define the displacement from the spring's rest length as $u_n \equiv x_{n+1} - x_n - a$, such that (1) becomes

$$\begin{aligned}\beta E &= \frac{K}{2}\sum_{n=0}^{N-1} u_n^2 + \frac{K'}{2}\sum_{n=0}^{N-1}(u_{n+1} + u_n)^2 \\ &= \frac{K + 2K'}{2}\sum_{n=0}^{N-1} u_n^2 + K'\sum_{n=0}^{N-1} u_n u_{n+1}\end{aligned} \tag{3}$$

(the above boundary conditions correspond to $u_N = u_0$). We introduce the discrete Fourier transform of the displacements

$$\mathcal{U}_q = \sum_{n=0}^{N-1} e^{-2\pi i q n/N}\, u_n, \tag{4}$$

with $q = -(N-1)/2, -(N-3)/2, \ldots (N-1)/2$ (assuming N odd) referred to as momentum here. Accordingly, the inverse Fourier transform is given by

$$u_n = \frac{1}{N}\sum_q e^{2\pi i q n/N}\, \mathcal{U}_q, \tag{5}$$

where the sum runs over the above given values of q. Since the u_n are real variables, we have $\mathcal{U}_q^* = \mathcal{U}_{-q}$. In momentum space, the energy then becomes

$$\beta E = \frac{1}{2N}\sum_q \widetilde{K}_q |\mathcal{U}_q|^2. \tag{6}$$

The stiffness of the mode with momentum q is

$$\widetilde{K}_q \equiv K + 4K' \cos^2\frac{\pi q}{N}. \tag{7}$$

Figure 1(b) shows $\widetilde{K}_q$ for $K = 1$ and $K' = 3$. The stability condition of the system $\widetilde{K}_q > 0$ for all q requires $K > 0$ and $K' > -K/4$, allowing for both positive and negative K'. To illustrate the meaning of these two cases, we note that K' enters as a diagonal coupling and as a cross term of the form $K' u_{n+1} u_n$, see (3). If $K' > 0$, configurations with u_n and u_{n+1} of opposite sign are favored, which leads to an anticorrelation of neighboring u_n. Conversely, negative K' leads to a positive correlation of neighboring u_n. It turns out that for DNA only the former case is observed.

The equipartition theorem, applied to (6), gives

$$\langle \mathcal{U}_q \mathcal{U}_{q'} \rangle = N \widetilde{K}_q^{-1} \, \delta_{q,-q'}, \tag{8}$$

where $\delta_{n,k}$ is the Kronecker delta. Moreover, collective m-step fluctuations can be expressed as

$$x_m - x_0 - ma = \sum_{n=0}^{m-1} u_n = \frac{1}{N} \sum_q \frac{\sin \frac{\pi q m}{N}}{\sin \frac{\pi q}{N}} e^{i \pi q (m-1)/N} \mathcal{U}_q. \tag{9}$$

Combining (2), (8), and (9) we find

$$\frac{m}{K_m} = \frac{1}{N^2} \sum_q \frac{\sin^2 \frac{\pi q m}{N}}{\sin^2 \frac{\pi q}{N}} \langle |\mathcal{U}_q|^2 \rangle = \frac{1}{N} \sum_q \frac{\sin^2 \frac{\pi q m}{N}}{\widetilde{K}_q \sin^2 \frac{\pi q}{N}}. \tag{10}$$

In the limit $N \to \infty$, one can replace the discrete sum with an integral

$$\frac{1}{K_m} = \frac{1}{m\pi} \int_{-\pi/2}^{\pi/2} \frac{\sin^2 my}{\sin^2 y} \frac{dy}{K + 4K' \cos^2 y}, \tag{11}$$

where we defined $y \equiv \pi q / N$ and used (7). For $m = 1$ and $m = 2$, a straightforward calculation shows that

$$K_1 = \sqrt{K(K + 4K')}, \tag{12}$$

$$K_2 = \frac{2K' \sqrt{K + 4K'}}{\sqrt{K + 4K'} - \sqrt{K}}. \tag{13}$$

In the asymptotic limit of large m, the factor $\sin^2(my)/\sin^2 y$ in (11) becomes increasingly peaked around $y = 0$. In the case $m \gg 1$, we find (see Appendix A)

$$K_m = \widetilde{K}_0 - \frac{2K'}{m} \sqrt{\frac{K + 4K'}{K}} + O\left(\frac{1}{m^2}\right), \tag{14}$$

Equations (12)–(14) show that the stiffness of the chain depends on the length scale at which fluctuations are observed. In the case $K' > 0$, one finds $K_1 < K_2 < \cdots < K_\infty$, e.g., the chain becomes increasingly stiffer at longer length scales [Fig. 1(b)]. The behavior is the opposite if $K' < 0$: the chain is softer at longer distances $K_1 > K_2 > \cdots > K_\infty$. As m increases, the contribution of large momenta to K_m gradually diminishes, until finally only the zero-momentum component ($q = 0$) contributes to the asymptotic stiffness $K_\infty = \widetilde{K}_0$. In the opposite limit ($m = 1$), K_1 becomes the harmonic mean of the momentum domain stiffnesses $\widetilde{K}_q$. Recall that the harmonic mean of N numbers ω_i with $i = 1, 2 \ldots N$ is defined as

$$\langle \omega \rangle_h = \left(\frac{1}{N} \sum_i \frac{1}{\omega_i} \right)^{-1}. \tag{15}$$

We consider now the effect of a local perturbation stretching one of the springs (say u_0), as illustrated in Fig. 2. This can be achieved by imposing a local force $f > 0$ on the selected degree of freedom such that the energy becomes

$$\beta E_f = \beta E - \beta f u_0 = \frac{1}{2N} \sum_q \widetilde{K}_q |\mathcal{U}_q|^2 - \frac{\beta f}{N} \sum_q \mathcal{U}_q, \tag{16}$$

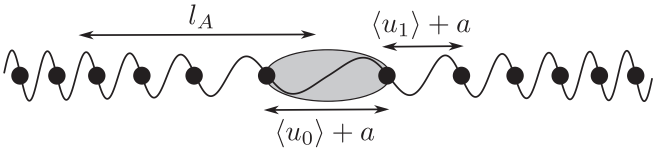

FIG. 2. Schematic illustration of the effect of a local perturbation at site $n = 0$ resulting in an exponentially decaying stretching profile $\langle u_m \rangle$, see Eq. (18). This depiction represents the case $K' < 0$, where the stretching decays monotonically (for the sake of clarity, we do not show next-neighbor springs).

with βE the unperturbed energy (6). The force stretches all modes to a nonzero average:

$$\langle \mathcal{U}_q \rangle = \frac{\beta f}{\widetilde{K}_q}. \tag{17}$$

The inverse Fourier transform then gives (for details, see Appendix (B1))

$$\langle u_m \rangle = \frac{\beta f}{N} \sum_q \frac{e^{2iqm/N}}{\widetilde{K}_q} = \frac{\beta f}{\pi} \int_{-\pi/2}^{\pi/2} \frac{e^{2iym} \, dy}{K + 4K' \cos^2 y}$$
$$= \frac{\beta f}{K_1} [-\text{sgn}(K')]^m \, e^{-m/l_A}, \tag{18}$$

with $m > 0$, sgn denoting the signum function and

$$\frac{1}{l_A} = -\ln \frac{|K_2 - K_1|}{K_2}. \tag{19}$$

Here K_1 and K_2 are the one-step and two-step stiffnesses defined in (12) and (13). We note that for $m = 0$, we get from (18) $K_1 \langle u_0 \rangle = \beta f$, showing again that K_1 is the stretching stiffness between neighboring sites. If $K' > 0$, the quantity $\langle u_m \rangle$ has an oscillatory decay, which can be easily understood from the coupling term $K' u_n u_{n+1}$ that contributes negatively if neighboring u_n have opposite signs. The same reasoning explains the monotonic decay if $K' < 0$. Note that in absence of length-scale dependence, which means that K_m does not depend on m, one has $l_A = 0$. Hence, in that case, a local perturbation does not affect flanking springs.

To conclude the analysis of the model, we remark that while our discussion here was limited to interactions ranging to next-nearest neighbors, i.e., involving just two spring constants (K and K'), the same formalism is directly applicable to systems involving further ranging interactions. In that case, (10) and (18) remain valid, but $\widetilde{K}_q$ will assume a more complicated form. We note a difference in the asymptotic behavior of the two quantities here analyzed. The m-dependent stiffness constant K_m approaches its asymptotic limit as $1/m$, see (14). This conclusion remains valid for an arbitrary $\widetilde{K}_q$ and is a consequence of the fact that K_m is determined by the average of $(\sum_{i=1}^m u_i)^2$, e.g., the stretching fluctuations over the whole segment. A local perturbation produces a shift in $\langle u_m \rangle$ which decays exponentially from the perturbation site (18), hence much more rapidly than K_m.

III. DNA ELASTICITY IN MOMENTUM SPACE

In our coarse-grained description of DNA, any configuration of a molecule consisting of $N + 1$ base pairs is fully described by a set of $N + 1$ orthonormal triads $\widehat{\mathcal{T}}_n = (\widehat{\mathbf{f}}_n \widehat{\mathbf{v}}_n \widehat{\mathbf{u}}_n)$,

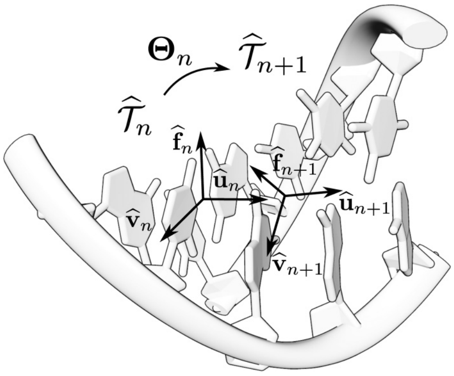

FIG. 3. Mapping of a DNA configuration into a rigid base-pair representation [23] that consists of a series of triads each attached to a single basepair, capturing the local geometry of the molecule. These triads are constructed from a set of three mutually orthogonal unit vectors $\widehat{\mathcal{T}}_n = (\widehat{\mathbf{f}}_n \widehat{\mathbf{v}}_n \widehat{\mathbf{u}}_n)$, where $\widehat{\mathbf{u}}_n$ is the local tangent, $\widehat{\mathbf{v}}_n$ connects the two backbones and $\widehat{\mathbf{f}}_n$ points toward the major groove. Deformations of the chain are parametrized by the rotation vectors $\boldsymbol{\Theta}_n$ rotating the triads $\widehat{\mathcal{T}}_n$ into their sequentially adjacent triads $\widehat{\mathcal{T}}_{n+1}$.

where $\widehat{\mathbf{f}}_n$, $\widehat{\mathbf{v}}_n$, and $\widehat{\mathbf{u}}_n$ are unit vectors capturing the local geometry of the base pair. We define $\widehat{\mathbf{u}}_n$ to be the local tangent and $\widehat{\mathbf{v}}_n$ to connect the two oppositely running backbones such that the remaining vector $\widehat{\mathbf{f}}_n = \widehat{\mathbf{v}}_n \times \widehat{\mathbf{u}}_n$ points toward the major groove [in the literature this frame is often indicated as $(\widehat{\mathbf{e}}_1 \widehat{\mathbf{e}}_2 \widehat{\mathbf{e}}_3)$ [16,26], here we use a different notation to avoid double indexing $\widehat{\mathbf{e}}_{1,n}$]. The spacial configuration of the molecule is given by the set of points connected by the vectors $a\widehat{\mathbf{u}}_n$, where a is the distance between consecutive base pairs. We assume this distance to be the constant value $a = 0.34$ nm. For simplicity, this description ignores stretching deformations. However, such could easily be included by replacing the connection vector $a\widehat{\mathbf{u}}_n$ by a variable three-component vector.

Up to a global rotation, a particular chain configuration is fully captured by the set of rotations that map each triad onto its consecutive triad, as illustrated in Fig. 3. It is convenient to parametrize these rotations by the corresponding Euler vectors $\boldsymbol{\Theta}$, i.e., the vectors parallel to the rotation axis with magnitude $\Theta = |\boldsymbol{\Theta}|$ equal to the rotation angle. To link the vector components to the local geometry, we express it in the basis of the local material frame:

$$\boldsymbol{\Theta}_n = a\tau_n \widehat{\mathbf{f}}_n + a\rho_n \widehat{\mathbf{v}}_n + a(\Omega_n + \omega_0)\widehat{\mathbf{u}}_n. \tag{20}$$

The components τ and ρ denote the two bending modes commonly referred to as tilt and roll [27], quantifying local bending over the axes $\widehat{\mathbf{f}}_n$ and $\widehat{\mathbf{v}}_n$, respectively. The total twist $\Omega_n + \omega_0$ (rotation around $\widehat{\mathbf{u}}_n$) has two components: Ω_n is the excess twist and $\omega_0 = 1.75\,\mathrm{nm}^{-1}$ the intrinsic twist of the double helix, corresponding to one turn of the helix every 10.5 base pairs. The deformation densities τ_n, ρ_n, and Ω_n of (20) have the dimension of inverse lengths and are expressed in nm^{-1}, while $a\tau_n$, $a\rho_n$, and $a\Omega_n$ are dimensionless and express rotation angles in radians.

The configuration $\tau_n = \rho_n = \Omega_n = 0$ (for all n) corresponds to a straight twisted rod with intrinsic twist ω_0, which is assumed to be the ground state of the system. Any deformation away from this state will be associated with a certain free energy. Expanding this free energy to lowest nonvanishing order around the ground state then corresponds to a regime of linear elasticity. In this paper, we limit our discussion to this regime. It is customary to describe DNA elasticity using on-site models, e.g., without interactions between neighboring sites. For instance, the Marko-Siggia model [26] is defined as

$$\beta E = \frac{a}{2} \sum_n \left(A^t \tau_n^2 + A^r \rho_n^2 + C\Omega_n^2 + 2G\rho_n\Omega_n \right), \tag{21}$$

where A^t, A^r, C, and G are stiffness parameters (we neglect in this description sequence-dependent effects and use constant stiffnesses). Besides the individual stiffnesses of tilt (A^t), roll (A^r), and twist (C), the model (21) is characterized by a nonvanishing twist-roll coupling (G), as expected from the symmetry of the molecule [26]. The effects of this coupling in the conformations of a DNA molecule were discussed recently in Refs. [19,21,28].

We generalize the elastic model to allow for interactions between further neighbors, employing a matrix representation

$$\beta E = \frac{a}{2} \sum_n \sum_m \boldsymbol{\Delta}_n^{\intercal} M_m \boldsymbol{\Delta}_{n+m}, \tag{22}$$

with $\boldsymbol{\Delta}_n^{\intercal} = (\tau_n, \rho_n, \Omega_n)$, and where the M_m are 3×3 matrices describing the couplings between sites separated by m steps. Stability of the model requires the on-site matrices M_0 to be positive definite.

Furthermore, for homogeneous directionally invariant chains, the general form of the matrices M_m can be deduced from symmetry considerations. Let us consider a configuration of $N + 1$ triads generated by N rotations parametrized by the vectors $\boldsymbol{\Theta}_n$ as in (20). The shape of the molecule remains unchanged for N rotations parametrized by vectors $\boldsymbol{\Theta}'_n \equiv -a\tau_n \widehat{\mathbf{f}}_n + a\rho_n \widehat{\mathbf{v}}_n + a(\Omega_n + \omega_0)\widehat{\mathbf{u}}_n$ applied in reversed order. This mapping of $\boldsymbol{\Theta}_n$ into $\boldsymbol{\Theta}'_n$ corresponds to a 180° rotation of the triad $\widehat{\mathcal{T}}_n$ around $\widehat{\mathbf{f}}_n$ and swaps the two strands [26]. Since such a coordinate transformation cannot change the energy, we see that for every m,

$$\boldsymbol{\Delta}_n^{\intercal} M_m \boldsymbol{\Delta}_{n+m} = \boldsymbol{\Delta}'^{\intercal}_{n+m} M_m \boldsymbol{\Delta}'_n, \tag{23}$$

with $\boldsymbol{\Delta}'_n{}^{\intercal} = (-\tau_n, \rho_n, \Omega_n)$. This implies that all off-diagonal terms in M_m involving τ have to be antisymmetric, while the remaining coupling (between ρ and Ω) is required to be symmetric. Hence, for homogeneous chains, the most general form of the matrices M_m is

$$M_m = \begin{pmatrix} A_m^t & A_m^{tr} & B_m \\ -A_m^{tr} & A_m^r & G_m \\ -B_m & G_m & C_m \end{pmatrix}. \tag{24}$$

This symmetry consideration implies for homogeneous on-site models (characterized by a single nonvanishing stiffness matrix M_0) that $A_0^{tr} = B_0 = 0$, which is the symmetry described by Marko and Siggia [26]. This is a consequence of M_0 being symmetric by construction while simultaneously satisfying (24).

We can rewrite the model (22) in momentum space as

$$\beta E = \frac{a}{2N} \sum_q \widetilde{\boldsymbol{\Delta}}_q^{\dagger} \widetilde{M}_q \widetilde{\boldsymbol{\Delta}}_q, \tag{25}$$

where $\widetilde{\boldsymbol{\Delta}}_q$ and $\widetilde{M}_q$ are the Fourier transform of $\boldsymbol{\Delta}_n$ and M_m, respectively, and † indicates the conjugate transpose. Stability of the model requires each of the Hermitian [29] matrices $\widetilde{M}_q$ to be positive definite, i.e., that all eigenvalues are positive. As indicated in (24), the matrices M_m may contain symmetric and antisymmetric components. Fourier transformation in m of the matrices (24) gives

$$\widetilde{M}_q = \begin{pmatrix} \widetilde{A}^t_q & i\widetilde{A}^{tr}_q & i\widetilde{B}_q \\ -i\widetilde{A}^{tr}_q & \widetilde{A}^r_q & \widetilde{G}_q \\ -i\widetilde{B}_q & \widetilde{G}_q & \widetilde{C}_q \end{pmatrix}, \tag{26}$$

where all entries $\widetilde{A}^t_q$, $\widetilde{A}^r_q$, $\widetilde{A}^{tr}_q$, $\widetilde{B}_q$, $\widetilde{C}_q$, and $\widetilde{G}_q$ are real variables. The off-diagonal terms $\widetilde{A}^{tr}_q$, $\widetilde{B}_q$ are odd functions of q (e.g., $\widetilde{A}^{tr}_{-q} = -\widetilde{A}^{tr}_q$), while all other terms are even functions of q.

The advantage of the momentum space representation is that modes with different q are independent (except for the coupling between q and $-q$). Strictly speaking, this is valid only if periodic boundary conditions are imposed such that full translational invariance is achieved. In absence of that, some boundary terms will appear, which, however, will be negligible for sufficiently large N.

Given an ensemble of deformation vectors $\boldsymbol{\Delta}_n$ the stiffness matrices can be obtained from the relation [30]

$$\langle \widetilde{\boldsymbol{\Delta}}_q \widetilde{\boldsymbol{\Delta}}^\dagger_q \rangle = \frac{N}{a} \widetilde{M}^{-1}_q, \tag{27}$$

where the 3×3 covariance matrix $\langle \widetilde{\boldsymbol{\Delta}}_q \widetilde{\boldsymbol{\Delta}}^\dagger_q \rangle$ is constructed from the ensemble averages of the products of the three components of the vector $\widetilde{\boldsymbol{\Delta}}^\intercal_q = (\widetilde{\tau}_q, \widetilde{\rho}_q, \widetilde{\Omega}_q)$. In the remainder of this section, we discuss the consequences of this model extension on various DNA properties: length dependence of persistence lengths and decays of local perturbations.

A. Twist persistence length

The twist-correlation function is defined as

$$\mathcal{C}_\mathrm{T}(m) = \left\langle \cos\left(a \sum_{n=0}^{m-1} \Omega_n \right) \right\rangle = \mathrm{Re}\langle e^{\,ia\sum_{n=0}^{m-1}\Omega_n} \rangle, \tag{28}$$

where Re denotes the real part. We are interested in the twist persistence length, which is the characteristic decay length of twist correlations

$$\frac{1}{l_\mathrm{T}} = -\frac{1}{ma} \ln \mathcal{C}_\mathrm{T}(m). \tag{29}$$

At this point, we present only a sketch of the calculation, as it is totally analogous to that of the elastic chain example discussed in detail in Sec. II. In like manner, we rewrite the sum in (28) in momentum space using expression (9). The variables $\widetilde{\Omega}_q$ for different momenta are independent, hence the total average (28) factorizes in terms of the form $\langle \exp(i\alpha_q \widetilde{\Omega}_q + i\alpha_{-q}\widetilde{\Omega}_{-q}) \rangle$ (it is convenient to group terms q and $-q$ together). Using the property of Gaussian variables

$$\langle e^{\pm i\alpha X} \rangle = e^{-\frac{\alpha^2}{2}\langle X^2 \rangle}, \tag{30}$$

we obtain in the limit $N \to \infty$,

$$\frac{1}{l_\mathrm{T}} = \frac{a}{2\pi m} \int_{-\pi/2}^{\pi/2} \frac{\sin^2 my}{\sin^2 y} \frac{\langle |\widetilde{\Omega}_q|^2 \rangle}{N}\, dy, \tag{31}$$

which is analogous to (11) and where we again used $y \equiv \pi q/N$. Just as in the example of Sec. II, the integral is dominated by smaller and smaller y contributions as m increases. The asymptotic twist persistence length ($m \to \infty$) is finally entirely governed by the zero-momentum component:

$$\frac{1}{l_\mathrm{T}} = \frac{a}{2N} \langle \widetilde{\Omega}^2_0 \rangle. \tag{32}$$

B. Bending persistence length

From the tangent-tangent correlation function

$$\mathcal{C}_\mathrm{B}(m) = \langle \widehat{\mathbf{u}}_0 \cdot \widehat{\mathbf{u}}_m \rangle, \tag{33}$$

one obtains the bending persistence length:

$$\frac{1}{l_\mathrm{B}} = -\frac{1}{ma} \ln \mathcal{C}_\mathrm{B}(m). \tag{34}$$

The twist-correlation function could be expressed exactly in terms of the deformation vectors $\boldsymbol{\Delta}_n$. However, establishing such a connection for $\mathcal{C}_\mathrm{B}$ requires some approximations. Under the assumption that the rotations connecting neighboring triads are dominated by the intrinsic twist component ω_0, we derived the following expression for the bending persistence length (for details, see Appendix C):

$$\frac{1}{l_\mathrm{B}} = \frac{a}{\pi m} \int_{-\pi/2}^{\pi/2} \frac{\sin^2 my}{\sin^2 y} \frac{\Psi_{q+\Delta q} + \Psi_{q-\Delta q}}{N}\, dy, \tag{35}$$

where we defined $\Delta q \equiv Na\omega_0/(2\pi)$ and

$$\Psi_q \equiv \frac{1-\cos(a\omega_0)}{2(a\omega_0)^2} \langle |\widetilde{\tau}_q|^2 + |\widetilde{\rho}_q|^2 \rangle. \tag{36}$$

This relation resembles Eq. (31) with the difference that here the $y(q)$ contributions of the momentum space bending deformations (tilt and roll) are replaced by the mean of the shifted momenta $q \pm \Delta q$. This stems from the fact that to appropriately connect local bending deformations to the total deformation of a given multistep segment (say from $\widehat{\mathbf{u}}_0$ to $\widehat{\mathbf{u}}_m$), one needs to rotate the local reference frames to unwind the intrinsic helical twist. Δq is indeed the momentum shift associated with the DNA intrinsic twist. As we integrate in the rescaled variable $y = \pi q/N$, the momentum shift corresponds to $\Delta y = a\omega_0/2 \approx \pi/10.5$, e.g., approximately one-tenth of the y domain (10.5 is the number of base pairs for a full turn of the double helix). In the limit $m \to \infty$, the $q = y = 0$ term is selected from the integral, and the asymptotic persistence length becomes [using (A3)]

$$\frac{1}{l_\mathrm{B}} = \frac{1-\cos(a\omega_0)}{a\omega_0^2 N} \langle |\widetilde{\tau}_{\Delta q}|^2 + |\widetilde{\rho}_{\Delta q}|^2 \rangle. \tag{37}$$

C. Local perturbations

Repeating the procedure applied to the linear chain model of Sec. II, we add a local perturbation at a given site of the DNA. This perturbation is introduced by means of generalized forces acting on the rotational degrees of freedom associated with that site —again we choose the site $n = 0$ but translational invariance implies that the results are equally valid for any given site—so the energy becomes

$$\begin{aligned}\beta E_{\mathbf{f}} &= \beta E - \beta \mathbf{f}^{\intercal} \mathbf{\Delta}_0 \\ &= \frac{a}{2N} \sum_q \left(\widetilde{\mathbf{\Delta}}_q^{\intercal} - \frac{\beta}{a} \mathbf{f}^{\intercal} \widetilde{M}_q^{-1} \right) \widetilde{M}_q \left(\widetilde{\mathbf{\Delta}}_q - \widetilde{M}_q^{-1} \frac{\beta}{a} \mathbf{f} \right) \\ &\quad - \frac{\beta^2}{2Na} \mathbf{f}^{\intercal} \widetilde{M}_q^{-1} \mathbf{f}, \end{aligned} \tag{38}$$

where βE is the unperturbed energy (25) and $\mathbf{\Delta}_0^{\intercal} = (\tau_0, \rho_0, \Omega_0)$. The vector $\mathbf{f}^{\intercal} = (f_\tau, f_\rho, f_\Omega)$ contains three components coupling to tilt, roll, and twist, respectively. These generalized forces shift the average $\widetilde{\mathbf{\Delta}}_q$ to the nonzero value

$$\langle \widetilde{\mathbf{\Delta}}_q^{\intercal} \rangle = \frac{\beta}{a} \mathbf{f}^{\intercal} \widetilde{M}_q^{-1}, \tag{39}$$

which is the equivalent of (17). In the DNA case, the calculation involves the inversion of the 3×3 matrix $\widetilde{M}_q$,

$$\widetilde{M}_q^{-1} = \frac{\mathrm{Adj}[\widetilde{M}_q]}{\det \widetilde{M}_q}, \tag{40}$$

where Adj[.] denotes the adjoint matrix. Combining (39) and (40) and performing the inverse Fourier transform, we obtain

$$\langle \mathbf{\Delta}_m^{\intercal} \rangle = \frac{\beta}{\pi} \int_{-\pi/2}^{\pi/2} \frac{\mathbf{f}^{\intercal}\, \mathrm{Adj}[\widetilde{M}_q]}{\det \widetilde{M}_q} e^{2iym}\, dy, \tag{41}$$

which is analogous to Eq. (18), derived for the toy model. As in that case, Eq. (41) gives rise to an exponential decay for large m: $\langle \mathbf{\Delta}_m \rangle \sim \exp(-ma/l_{\mathrm{A}})$. The characteristic decay length l_{A} is given by the poles closest to the real axis of the integrand (see Appendix (B1)). We note that stability of the energy (25) requires $\det \widetilde{M}_q > 0$ in the real q domain. Hence poles necessarily have an imaginary component responsible for the exponential decay. In practice, this integral can be evaluated numerically from empirically obtained $\widetilde{M}_q$.

IV. DNA ELASTICITY IN COARSE-GRAINED AND ALL-ATOM MODELS

We discuss and compare here the elasticity of the coarse-grained DNA model oxDNA [11] and of an all atom model. The main focus is the calculation of $\widetilde{M}_q$ from which various quantities are obtained, following the framework discussed in the previous section.

A. oxDNA

The oxDNA model treats nucleotides as single rigid objects that mutually interact via multiple sites representing the most significant interbase interactions: backbone connectivity, base pairing, and base stacking. These interactions are parametrized so as to reproduce thermodynamical, structural, and mechanical properties of DNA [11]. oxDNA has been used to study a broad range of processes such as DNA melting, hybridization, supercoiling, looping, DNA strand-displacement mechanisms, DNA gels, nanotubes, and origami [31–37]. Here we focus exclusively on oxDNA2 [38], a version of the model with asymmetric major and minor grooves. We used the procedure outlined in Ref. [16] to map the oxDNA coordinates to orthonormal triads $(\widehat{\mathbf{f}}_n \widehat{\mathbf{v}}_n \widehat{\mathbf{u}}_n)$ (Fig. 3). This mapping is not unique and a few alternative definitions have been discussed in Ref. [16]. Differences in triads are carried over to the rotational modes $\mathbf{\Delta}_n$, which leads to slightly different elastic behavior. However, we observe the Fourier spectra of the couplings to exhibit the same general features. In particular, alternative triads give the same behavior at small q (same asymptotic elasticity) and follow the same trend from small to large q behavior. We will present here the results from triad2, as defined in Ref. [16].

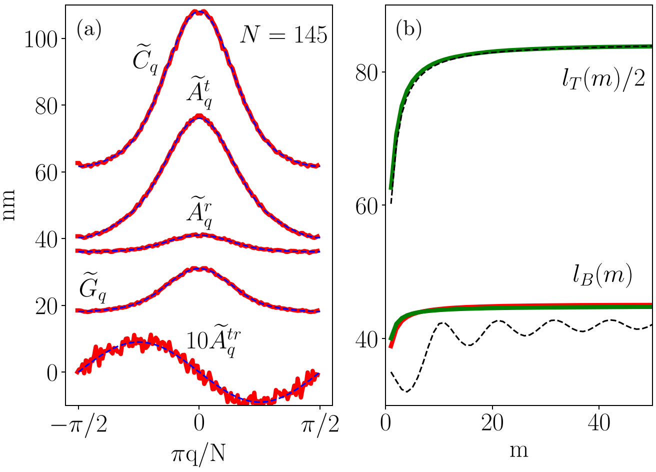

FIG. 4. (a) Red dots: Simulation data reporting the entries of the stiffness matrix in momentum space $\widetilde{M}_q$ for oxDNA2 as obtained from Eq. (27) for a sequence of length 150. In the analysis, two nucleotides at the two ends were eliminated, which gives 146 triads and thus $N = 145$ deformation vectors $\mathbf{\Delta}_m$. The units are in nm. The entry $\widetilde{A}_q^{tr}$ has been multiplied by a factor of 10 to facilitate its visibility. The stiffness matrix has the structure given in (26). All its entries are symmetric in q, except for the tilt-roll term $\widetilde{A}_q^{tr}$ which is antisymmetric. Blue dashed lines: Fits of the data to Eqs. (42) and (43), with fitting parameters given in Table I. (b) Plots of l_{B} and $l_{\mathrm{T}}/2$ vs m the relative distance in numbers of base-pair steps between the considered segments. Green lines are obtained from the stiffness matrix data using Eqs. (31) and (35). The red line is the approximation (C19). In this case, the difference between the two approximations for l_{B} is very small. Black dashed lines are obtained by direct calculations of correlation functions from simulations. The oscillatory behavior of the bending persistence length stems from a light helicity of the traced contour.

Using molecular dynamics trajectories of oxDNA2 (details about simulations can be found in Ref. [16]), we computed the Fourier spectra of the rotational deformations $\widetilde{\mathbf{\Delta}}_q^{\intercal} = (\widetilde{\tau}_q, \widetilde{\rho}_q, \widetilde{\Omega}_q)$. The stiffness matrices $\widetilde{M}_q$ were then obtained by utilizing Eq. (27). The matrix entries versus rescaled momentum $y \equiv \pi q/N$ are plotted in Fig. 4(a). These matrices indeed follow the structure (26) as predicted by the symmetry consideration. The antisymmetric components turn out to be very small, with $\widetilde{B}_q$ virtually zero. The only significant off-diagonal term in oxDNA2 is the twist-roll coupling $\widetilde{G}_q$ [16]. We note that $\widetilde{A}_q^r$, the roll stiffness, is very weakly dependent on q as compared to the other entries. This weak dependence indicates that the roll-roll interaction is dominated by the on-site term ρ_n^2. The strong dependence on q for tilt-tilt and twist-twist terms implies significant contributions from off-site interactions $\tau_n \tau_{n+m}$ and $\Omega_n \Omega_{n+m}$, with $m > 0$.

TABLE I. Summary of the stiffnesses in oxDNA2. X_m are the fitting coefficients used in Eqs. (42) and (43). The two rightmost columns give the stiffnesses at $q = 0$ and $q = \Delta q$, as representatives of the long length-scale behavior [see Eqs. (32) and (37)]. The last two lines give the persistence lengths as obtained from Eqs. (31) and (35). We give the local ($m = 1$) value and the asymptotic one ($m \to \infty$). All parameters are given in nm.

	X_0	X_1	X_2	X_3	$q = 0$	$q = \Delta q$
$\widetilde{A}^t_q$	54	17	4.0	1.1	76	69
$\widetilde{A}^r_q$	38	2	0.8	0.2	41	40
$\widetilde{C}_q$	78	22	6.5	1.3	108	98
$\widetilde{G}_q$	23	6.0	1.9	0.4	31	28
$\widetilde{A}^{tr}_q$		-0.9			0	-0.5
l_B		40 ($m = 1$)		45 ($m \to \infty$)		
$l_T/2$		63 ($m = 1$)		84 ($m \to \infty$)		

To quantify these effects, the inverse Fourier transform of the data in Fig. 4(a) was computed so as to obtain the couplings in real space [39]. The Fourier series of the elements of the stiffness matrix which are even or odd in q are given by

$$\widetilde{X}_q^{\text{even}} = \sum_m X_m \cos \frac{2m\pi q}{N}, \tag{42}$$

$$\widetilde{X}_q^{\text{odd}} = \sum_m X_m \sin \frac{2m\pi q}{N}, \tag{43}$$

where X_m is the real-space stiffness associated to couplings between sites n and $n + m$ [40].

For the even terms, we truncated the series to the first four components, while in view of the uncertainties of the small odd term $\widetilde{A}^{tr}_q$ we used a single term. The best fits to the data are shown as dashed blue lines in Fig. 4(a). Table I gives the values of the corresponding coefficients X_m resulting from the fits. The coefficients decrease rapidly with m, but there are significant off-site components for $\widetilde{C}_q$ and $\widetilde{A}^r_q$, reflecting the strong q-dependence observed in Fig. 4(a). Twist and bend fluctuations are linked to the elements of the stiffness matrix via the covariance matrix (27). Neglecting the small contribution of $\widetilde{A}^{tr}_q$ and inverting $\widetilde{M}_q$ we get

$$\frac{a\langle|\widetilde{\Omega}_q|^2\rangle}{N} = \frac{1}{\widetilde{C}_q - \widetilde{G}_q^2/\widetilde{A}^r_q} \tag{44}$$

and

$$\frac{a\langle|\widetilde{\tau}_q|^2 + |\widetilde{\rho}_q|^2\rangle}{N} = \frac{1}{\widetilde{A}^t_q} + \frac{1}{\widetilde{A}^r_q - \widetilde{G}_q^2/\widetilde{C}_q}, \tag{45}$$

Inserting (44) in (31), we can estimate the twist persistence length $l_T(m)$ from the stiffness data using the truncated Fourier series as numerical estimates for $\widetilde{A}^r_q$, $\widetilde{G}_q$, and $\widetilde{C}_q$. In a similar way, inserting (45) into Eq. (35) allows us to calculate the bending persistence length. The results of these calculations are shown in Fig. 4(b) as solid green lines. The red solid line is the approximation (C19) [41]. Dashed black lines show the direct calculations of the bending persistence length as deduced from the decay length of the respective correlation functions [(28) and (33)]. While there is excellent overlap between dashed and solid lines for l_T, some deviations of a few nm are visible in l_B. The overlap in l_T was expected as (31) is exact, while both expressions (C19) and (35) [red and green lines in Fig. 4(b)] involve approximations. Note also that l_B as deduced from the correlation function exhibits damped oscillatory behavior stemming from a light helicity of the used set of triads.

The last two lines of Table I give the local ($m = 1$) and asymptotic ($m \to \infty$) values of the persistence lengths as obtained from (31) and (35). Both bending and torsional persistence lengths are smaller at short distances as compared to their asymptotic values, however, the effect is modest for l_B, while much stronger length-dependent variability is observed in l_T. This can be understood from the elements of the stiffness matrix. Torsional persistence is primarily determined by $\widetilde{C}_q$ [Eq. (44)] which has a large q dependence, causing strong length-scale effects in l_T. On the other side, the bending stiffness is determined by the harmonic mean of tilt and rescaled roll stiffnesses (45), which is dominated by the softer roll component. The weak dependence of $\widetilde{A}^r_q$ on q in Fig. 4(a), indicating small off-site roll-roll couplings, is the cause of the modest length-scale dependence of l_B.

B. All atom

All-atom simulations of double stranded DNA of two different lengths were performed. Details of setup, force fields, methodology, and sequences used can be found in Appendix D. Tilt, roll, and twist variables were obtained from simulation data using an implementation of the algorithm underlying Curves+ [27]. Subtracting the averages, we obtained the excess values $\mathbf{\Delta}_n^{\intercal} = (\tau_n, \rho_n, \Omega_n)$. Local elasticity in all-atom models of DNA is dependent on the type of base pairs, as opposed to the homogeneous oxDNA model. Using the relation (27), we derived an effective stiffness matrix $\widetilde{M}_q$. The procedure builds up an equivalent homogeneous model which shares the same covariance matrix as the original data set by matching the second moments of the fluctuations in Fourier space. For a system breaking translational invariance, in general, the correlator $\langle \widetilde{\mathbf{\Delta}}_q \widetilde{\mathbf{\Delta}}_{q'}^{\dagger} \rangle$ is nonzero also for $q \neq q'$. In constructing the average stiffness matrix, we ignore these off-diagonal terms, which are expected to have weaker effect as the system size grows, where effective translation invariance is recovered.

Figure 5 shows the elements of $\widetilde{M}_q$ in function of $y = \pi q/N$ as obtained from this procedure (red dots and black squares). The lengths simulated correspond to (a) 20-mers and (b) 32-mers, averaged over ten and three different sequences, respectively. Two nucleotides at each end were removed from the analysis to mitigate end effects. Hence Fig. 5 shows the Fourier transforms on (a) $N = 15$ and (b) $N = 27$ data points. Despite the difference in length, the two sets exhibit quantitatively very similar stiffnesses. The data share several common features with the oxDNA simulations of Fig. 4: The tilt $\widetilde{A}^t_q$ and twist $\widetilde{C}_q$ stiffnesses are strongly q dependent, indicating considerable contributions from off-site interactions. Just as for oxDNA, the roll stiffness $\widetilde{A}^r_q$ depends very weakly on q and again the only symmetric off-diagonal term of the stiffness matrix is the twist-roll coupling $\widetilde{G}_q$. Contrasting the oxDNA results, in the all-atom data the tilt stiffness is larger than the twist stiffness $\widetilde{A}^t_q > \widetilde{C}_q$ and their values are quantitatively

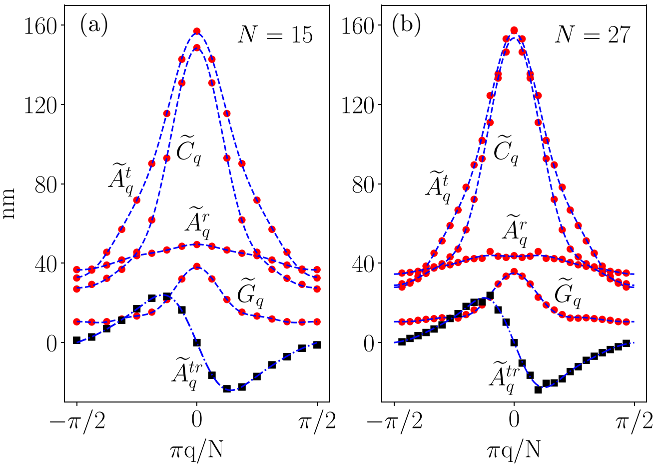

FIG. 5. Red dots and solid squares: Elements of the stiffness matrix $\widetilde{M}_q$ as obtained from all-atom data for sequences of length (a) $N = 20$ (average of 9 seq.) and (b) $N = 32$ (average of three seq.). Dashed lines: Fits of the forms (42) and (43).

much larger. In addition, the q-odd tilt-roll coupling $\widetilde{A}_q^{tr}$ is much more prominent than in oxDNA. Our results are consistent with other all-atom studies. Lankas *et al.* [2] observed that tilt and twist are highly anticorrelated beyond neighboring steps, while roll is very weakly correlated, which is in agreement with the present work. In addition, various degrees of correlations have been discussed by other authors [5,23,24].

Table II shows the results of the fits of the elements of $\widetilde{M}_q$ to Eqs. (42) and (43). The coefficients X_m decrease significantly with m but more gradually as compared to oxDNA, indicating more pronounced off-site interactions. Overall, there is only a small difference between the two data sets, which is indicative for weak finite-size effects. Using the coefficients X_m of the $N = 27$ data set as representatives for the couplings of a long DNA sequence, we invoked (31) and (35) to estimate the twist and bending persistence lengths. Results are shown in Fig. 6(a). As in oxDNA $l_{\rm T}$ has a strong length-scale dependence, while for $l_{\rm B}$ this dependence is much more modest. The variability of $l_{\rm T}$ across different length scales is much larger in the all-atom data than in oxDNA. This is due to the much stronger q dependence of the stiffnesses of the former as can be seen when comparing Fig. 5 to Fig. 4. Interestingly, $l_{\rm T}/2$ approaches an asymptotic value close to 130 nm, which is not far from the torsional stiffnesses (120 nm) measured in magnetic tweezers [42]. This technique probes the torsional elasticity by tracing the twist fluctuations of the ends of stretched DNA molecules of several kilobases length. The recent atomistic simulation study by Velasco-Berreleza *et al.* [8] found a similarly strong length dependence of the torsional fluctuations, although their asymptotic estimate indicates $l_{\rm T}/2 \approx 90$ nm. We note here that $l_{\rm T}$ at all length scales is not only determined by the twist stiffness $\widetilde{C}_q$, but also by other stiffnesses. In oxDNA, twist fluctuations are also influenced by $\widetilde{G}_q$ and $\widetilde{A}_q^{tr}$, see Eq. (44). The relation is even more elaborate if one includes the tilt-roll coupling $\widetilde{A}_q^{tr}$, which is non-negligible in all atom data.

TABLE II. All-atom data for 20-mers ($N = 15$) and 32-mers ($N = 27$) averaged over ten and three different oligomers, respectively. All parameters are given in nm.

$N = 15$	X_0	X_1	X_2	X_3	X_4	$q = 0$	$q = \Delta q$
$\widetilde{A}_q^t$	82	56	11	5.8	1.3	156	130
$\widetilde{A}_q^r$	43	5.9	−0.4	0.6	0.3	50	48
$\widetilde{C}_q$	65	52	21	8.5	1.9	148	112
$\widetilde{G}_q$	17	11	5.4	2.9	1.4	38	27
$\widetilde{A}_q^{tr}$		−19	−8.4	0.3	−0.6	0	−21
$N = 27$	X_0	X_1	X_2	X_3	X_4	$q = 0$	$q = \Delta q$
$\widetilde{A}_q^t$	75	57	14	6.7	2.6	156	125
$\widetilde{A}_q^r$	40	4.7	−0.4	−0.2	−0.5	43	44
$\widetilde{C}_q$	67	53	23	9.3	1.4	154	116
$\widetilde{G}_q$	17	9.3	5.0	2.9	1.0	35	25
$\widetilde{A}_q^{tr}$		−16	−8.9	0.4	−1.2	0	−21
$l_{\rm B}$		42 ($m = 1$)		61 ($m \to \infty$)			
$l_{\rm T}/2$		43 ($m = 1$)		125 ($m \to \infty$)			

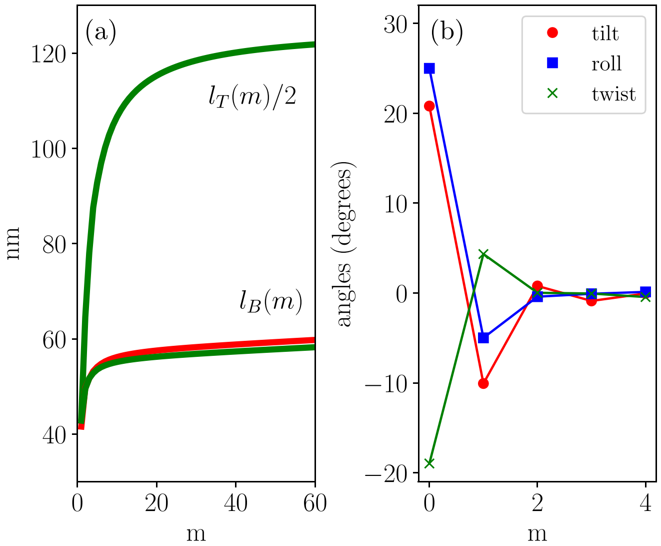

FIG. 6. (a) Estimated length-scale dependence of the persistence lengths as obtained from the analysis of the all-atom data in Table II. Assuming that these data are representatives for the behavior of very long sequences, we used Eqs. (35) and (31) to calculate $l_{\rm B}$ and $l_{\rm T}$ (green lines). The red line is the approximation (C19) for $l_{\rm B}$. (b) Calculation of the propagation of perturbations induced by generalized forces acting on the site $m = 0$. This data is calculated with Eq. (41) using the data in Table II. Results are given in degrees (the quantities plotted are $180\, a\tau/\pi$, $180\, a\rho/\pi$, and $180\, a\Omega/\pi$).

Note that in the context of atomistic simulations, the ground-state of the chain is generally not straight but exhibits intrinsic curvature brought forward by static bending components (i.e., the average values of tilt and roll are nonzero). These static bending components give rise to an additional contribution to the decorrelation of distal tangents on top of the dynamic decorrelation, generated by thermal fluctuations. Accordingly, the persistence length obtained from the decay of the tangent-tangent correlation functions is usually decomposed into the dynamic persistence length l_d and the static persistence length l_s [43]. To determine the characteristic decay-length of the tangent-tangent correlation function, one has to consider both effects. In the present paper, we average over sequence-dependent effects and omit an analysis of static bending contributions as this effect has been studied in detail in the prior literature [44]. Accordingly, the bending persistence length reported here should be interpreted as the dynamic persistence length. Mitchell *et al.* [44] recently estimated both the static and dynamic components of the persistence length using cgDNA, a coarse-grained rigid base model. They report an averaged dynamic persistence length of 58.8 nm, which is in very close agreement with the 61 nm predicted by our analysis (see Table II). The torsional persistence length, on the contrary, does not have a static component as it is defined from the sum of the excess of local twist densities, see Eq. (28).

Figure 6(b) shows our calculation of the response of a DNA molecule to a generalized force imposed on a certain base-pair step, as given by the integral (41). The generalized force $(f_\tau, f_\rho, f_\Omega)$ was tuned to shift the average deformations ($\langle\tau_0\rangle$, $\langle\rho_0\rangle$, and $\langle\Omega_0\rangle$) from zero to some finite angles (20°, 25°, and −20° for tilt, roll, and twist, respectively). Due to the presence of nonlocal couplings, neighboring steps are expected to also be effected by this imposed force. The calculation shows that the resulting shift in the average values decay very rapidly to zero, which is the unperturbed value, with angles being negligibly small already at $m = 2$. Although off-site couplings are capable of carrying the effect of a local perturbation to distant flanking sites, the characteristic decay length l_A is quite small. Why are the twist and, to a more limited extent, the bending elasticity varying so much with the length scale [Fig. 6(a)] while local perturbations [Fig. 6(b)] decay so rapidly? To understand this issue, it is useful to go back to the toy model of Sec. II. At different length scales, the elasticity is governed by different stiffnesses ranging from K_1 to K_∞, where the asymptotic value is approached as $1/m$ for large m [see Eq. (14)]. A local perturbation, on the contrary, decays exponentially with a length linked to the relative difference between the two local elastic constants K_1 and K_2, see Eq. (19).

V. DISCUSSION

In this paper, we investigated the effects of interactions in DNA models that extend beyond nearest neighbors (off-site couplings), as discussed in several prior works [5,23,24]. Our analysis is based on the calculation of the stiffness matrix in momentum space $\widetilde{M}_q$ for oxDNA and all-atom models. Both systems show very similar behavior, which is presumably a consequence of the geometrical structure of the double helix. The set of matrices $\widetilde{M}_q$ encodes both the asymptotic long length-scale stiffness $q = 0$ as well as the short scale behavior obtained by taking the harmonic mean of the matrices. We summarize here the main findings.

A. General structure of the coupling matrices

Both oxDNA and all-atom data indicate that the general form of the off-site coupling matrices can be understood from symmetry arguments, generalizing those used to describe on-site interactions [26]. This symmetry requires the functional form of homogeneous models to be invariant under reversal of the curvilinear coordinate, such that the first segment becomes the last and vice versa. The resulting generic form of $\widetilde{M}_q$ is given by Eq. (26) and contains terms which are either even or odd in q. As odd terms vanish in the limit $q \to 0$, they have a weak impact on the asymptotic length scale elasticity but they turn out to be more relevant at short length scales. Our analysis confirms the findings of previous studies [16], in showing that the twist-roll coupling $\widetilde{G}_q$ (even function of q) is the dominant off-diagonal stiffness coefficient.

B. Length dependence of persistence lengths

Our analysis has shown that of the three rotational modes, tilt (τ) and twist (Ω) exhibit significant off-site couplings. This can be seen from the strong q dependence of the respective momentum space couplings ($\widetilde{A}^t_q$ and $\widetilde{C}_q$) as shown in Figs. 4(a) and 5 or, equivalently, in the appreciable real space coupling that extend up the fourth neighbor in the case of the atomistic simulations (see Table II). On the other hand, the remaining mode roll (ρ) shows but modest off-site interactions, i.e., a very weak q dependence of the momentum space couplings ($\widetilde{A}^r_q$). In all cases, the mode stiffness is softer locally and becomes increasingly stiffer toward the asymptotic long-range regime. From the behavior of these three modes, one can understand the length dependence of the twist and bending persistence length. The twist persistence length l_T is fully determined by the behavior of the twist degree of freedom and therefore mirrors its strong length dependence [see Fig. 4(b)], which is in agreement with previous studies [5]. For oxDNA2 this manifests in about a 35% increase in stiffness from the local to the asymptotic elasticity. The bending persistence length l_B is determined by the harmonic mean of the stiffnesses governing the fluctuations of the two bending modes τ and ρ, which is dominated by the softer ρ mode [see Eqs. (37) and (45)]. Accordingly, the weak length dependence of this mode translates into a likewise behavior of the bending persistence length. We observed similar effects in the all-atom data, although the difference in torsional elasticity at short and very long length scales is much larger in that case, as illustrated in Fig. 6. This strong length-scale dependence of the torsional elasticity can potentially explain the divergence between estimates obtained with different experimental methods [8]. Studies that employ local probing methods find systematically lower stiffnesses as compared to studies in which larger length scales are considered, as is the case for magnetic tweezers (for a list of different estimates and methods used, see Supplemental Material of Ref. [45]).

C. Local perturbations

Our model predicts that local DNA deformations such as an imposed bending or twist angle at a given site induces structural changes of the flanking sites up to some characteristic distance. This distance depends both on the magnitude of the off-diagonal couplings and the range of the interactions. For the analyzed models, we find that the effect is rather modest, with the perturbation involving just three flanking sites. Experiments analyzing DNA-protein interactions have highlighted a few cases of distal allosteric effects [46,47], where the binding of a protein at a given site increases the binding affinity to a second protein. This distance is of about 15–20 nucleotides. A more common phenomenon is that of proximal allostery, which involves the binding of small molecules in the DNA minor groove altering the corresponding major groove binding site affinity for a protein (see, for example, the discussion in Refs. [48,49]). Our analysis indicates that, within linear elasticity, distal allostery is rather modest as compared to the distal effects seen in these experiments [46,47]. This short perturbation range was obtained from the average elastic behavior of the considered sequences. It remains to be seen if some specific sequences can exhibit a much more pronounced effect. Beyond that, it is likely that, to fully account for the experimentally observed allostery, one would need to go beyond linear elasticity, see, e.g., Ref. [50].

D. On the origin of off-site couplings

Our analysis indicates that there are off-site couplings in DNA, confirming prior all-atom simulation data [23,24]. Off-site couplings are found both in oxDNA as well as in all-atom simulations, although the effects are generally stronger in the latter as seen from the comparison in Tables I and II. In both cases, we map simulation data to a simplified model described by the three local configuration coordinates tilt, roll, and twist. This mapping corresponds to a coarsening procedure, a step which often induces more complex couplings than those of the original coordinates. Rigorous bottom-up coarse-grained methods give rise to many body potentials of mean force, see, e.g., Ref. [51]. Indeed, an analysis of molecular dynamics simulations of DNA indicated that the nonlocal couplings are due to the tracing out of some local degrees of freedom [23]. In the all-atom case, the solvent can play a role in inducing off-site interactions, as also suggested in Ref. [24]. In all-atom simulations, the couplings extend up to a maximum of four neighbors, corresponding to a distance of 1.2 nm along the DNA backbone. The Debye length of the solvent has a comparable value of $\lambda_D = 0.8$ nm (see Appendix D), suggesting that counterions can mediate an interaction over that range of distances.

ACKNOWLEDGMENTS

ES acknowledges financial support from Fonds Wetenschappelijk Onderzoek (FWO 1SB4219N). Part of this work was carried out on the Dutch national e-infrastructure with the support of SURF Cooperative (2020.015-Cartesius).

APPENDIX A: ASYMPTOTIC BEHAVIOR OF K_m

To estimate the asymptotic behavior of K_m, the elastic constant of the model introduced in Sec. II, as given in (11) we use the following identity:

$$\frac{\sin^2 my}{\sin^2 y}\frac{1}{K+4K'\cos^2 y} = \frac{\sin^2 my}{K+4K'\cos^2 y} + \frac{1}{4K'}\frac{\sin^2 my}{\sin^2 y}\left(1-\frac{K}{K+4K'\cos^2 y}\right). \tag{A1}$$

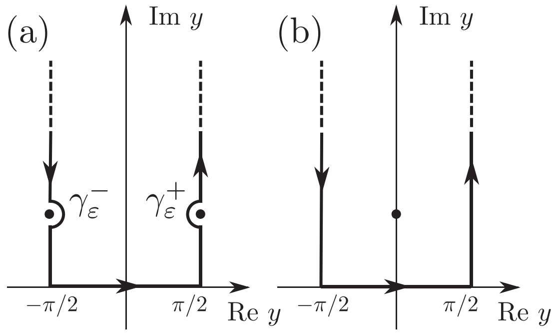

FIG. 7. Integration contours in the complex y plane used for the evaluation of the integral (B1). The two cases correspond to (a) $K' > 0$ and (b) $-K/4 < K' < 0$.

Integrating both sides and dividing by $m\pi$, we find

$$\frac{1}{K_m}\left(1+\frac{K}{4K'}\right) = \frac{1}{2mK_1} + \frac{1}{4K'} + O(e^{-\frac{m}{\lambda}}), \tag{A2}$$

where we used

$$\frac{1}{m\pi}\int_{-\pi/2}^{\pi/2}\frac{\sin^2 my}{\sin^2 y}\,dy = 1, \tag{A3}$$

and

$$\frac{1}{m\pi}\int_{-\pi/2}^{\pi/2}\frac{\sin^2 my\,dy}{K+4K'\cos^2 y} = \frac{1}{2mK_1} + O(e^{-\frac{m}{\lambda}}). \tag{A4}$$

$(e^{-m/\lambda})$ denotes a term exponentially small in m and of higher order with respect to the leading $1/m$ behavior. Inverting Eq. (A2) and substituting the value of K_1 of (12), one obtains the asymptotic expansion (14).

APPENDIX B: DECAY OF LOCAL PERTURBATION

We give here further details about the calculation of the integral in Eq. (18):

$$I = \frac{\beta f}{\pi}\int_{-\pi/2}^{\pi/2}\frac{e^{2iym}\,dy}{K+4K'\cos^2 y}. \tag{B1}$$

As mentioned earlier, stability of the model requires that either $K' > 0$ or $-K/4 < K' < 0$. We will discuss these two cases separately.

1. $K' > 0$

In this case, the integrand has two simple poles in $y = \pm\pi/2 + i\alpha$ with $\alpha > 0$ the solution of $\cosh^2\alpha = K/4K'$. We extend the integration over the contour indicated in Fig. 7(a), which is closed at infinity. The integral in this domain does not enclose any singularities hence it vanishes. The integrals along the two vertical lines cancel each other, due to symmetry, so one is left with

$$I + \frac{\beta f}{\pi}\int_{\gamma_\varepsilon^+\cup\gamma_\varepsilon^-}\frac{e^{2iym}\,dy}{K+4K'\cos^2 y} = 0, \tag{B2}$$

where

$$\gamma_\varepsilon^\pm(\phi) = i\alpha \pm \frac{\pi}{2} + \varepsilon e^{-i\phi} \tag{B3}$$

are the two small half-circles around the two poles. The integrations in these two domains pick up contributions from the poles and directly yield the expression (18). In particular, the oscillating behavior stems from the fact that the poles are in $\pm\pi/2$, which leads to the appearance of a factor $\exp(\pm im\pi) = (-1)^m$. The associated decay length is then simply given by $l_\mathrm{A} = 1/2\alpha$.

2. $-K/4 < K' < 0$

In this case, the integrand has a simple pole in $y = i\alpha$ with $\alpha > 0$ the solution of the equation $\cosh^2\alpha = K/4|K'|$. We extend the integration to the domain shown in Fig. 7(b). The integration picks up the residue from the pole along the imaginary axis. Thus, one can again obtain I. Note that, as the pole is purely imaginary, there are no oscillations but a pure exponential decay.

More complicated integrands will eventually contain several poles, giving rise to a sum of exponentials. The dominant contribution will be given by the pole in the semi-infinite strip $-\pi/2 \leqslant \mathrm{Re}(y) \leqslant \pi/2$, $\mathrm{Im}(z) > 0$, which is closest to the real axis.

APPENDIX C: BENDING PERSISTENCE LENGTH

The rotation operator mapping the triad $(\widehat{\mathbf{f}}_k\widehat{\mathbf{v}}_k\widehat{\mathbf{u}}_k)$ into $(\widehat{\mathbf{f}}_{k+1}\widehat{\mathbf{v}}_{k+1}\widehat{\mathbf{u}}_{k+1})$ can be expressed as

$$\mathcal{R}_k = \widehat{\mathbf{f}}_{k+1} \otimes \widehat{\mathbf{f}}_k + \widehat{\mathbf{v}}_{k+1} \otimes \widehat{\mathbf{v}}_k + \widehat{\mathbf{u}}_{k+1} \otimes \widehat{\mathbf{u}}_k. \tag{C1}$$

Here $\otimes$ denotes the tensor product, which transforms a generic vector $\mathbf{a}$ as follows:

$$(\mathbf{u} \otimes \mathbf{v})\mathbf{a} = (\mathbf{a} \cdot \mathbf{v})\mathbf{u}. \tag{C2}$$

From (C1), it follows that $\mathcal{R}_k\widehat{\mathbf{f}}_k = \widehat{\mathbf{f}}_{k+1}$, $\mathcal{R}_k\widehat{\mathbf{v}}_k = \widehat{\mathbf{v}}_{k+1}$ and $\mathcal{R}_k\widehat{\mathbf{u}}_k = \widehat{\mathbf{u}}_{k+1}$. An alternative axis-angle representation uses a unit vector $\widehat{\boldsymbol{\gamma}}$ as rotation axis and a rotation angle θ. For a counterclockwise rotation around $\widehat{\boldsymbol{\gamma}}$, this representation takes the form

$$\mathcal{R} = \cos\theta(1 - \widehat{\boldsymbol{\gamma}} \otimes \widehat{\boldsymbol{\gamma}}) + \sin\theta\,(\boldsymbol{\epsilon}\widehat{\boldsymbol{\gamma}}) + \widehat{\boldsymbol{\gamma}} \otimes \widehat{\boldsymbol{\gamma}}, \tag{C3}$$

where

$$(\epsilon\mathbf{u})\mathbf{a} = \mathbf{u} \times \mathbf{a}. \tag{C4}$$

One can easily verify from (C3) that $\mathcal{R}\widehat{\boldsymbol{\gamma}} = \widehat{\boldsymbol{\gamma}}$ and that for any unit vector $\widehat{\mathbf{a}}$ orthogonal to $\widehat{\boldsymbol{\gamma}}$ the following relations hold: (a) $\widehat{\boldsymbol{\gamma}} \cdot \mathcal{R}\widehat{\mathbf{a}} = 0$ and (b) $\widehat{\mathbf{a}} \cdot \mathcal{R}\widehat{\mathbf{a}} = \cos\theta$. This shows that the rotated vector $\mathcal{R}\widehat{\mathbf{a}}$ is orthogonal to the rotation axis and that it forms an angle θ with $\widehat{\mathbf{a}}$. As mentioned in the main text, tilt, roll, and twist are the components of the Euler vector with respect to the local triad

$$\boldsymbol{\Theta} = a\tau\widehat{\mathbf{f}} + a\rho\widehat{\mathbf{v}} + a(\Omega + \omega_0)\widehat{\mathbf{u}}, \tag{C5}$$

where its length $\Theta \equiv |\boldsymbol{\Theta}|$ gives the rotation angle. It is convenient to define

$$\mathrm{t} \equiv a\tau/\Theta, \qquad \mathrm{r} \equiv a\rho/\Theta, \qquad \mathrm{w} \equiv a(\Omega + \omega_0)/\Theta, \tag{C6}$$

for which $\mathrm{t}^2 + \mathrm{r}^2 + \mathrm{w}^2 = 1$ holds. Using (C3) with $\widehat{\boldsymbol{\gamma}} = \boldsymbol{\Theta}_k/\Theta_k$ and $\theta = \Theta_k$ and (C5) one finds

$$\begin{aligned}\widehat{\mathbf{u}}_{k+1} = \mathcal{R}_k\widehat{\mathbf{u}}_k &= \big[\cos\Theta_k + (1 - \cos\Theta_k)\mathrm{w}_k^2\big]\widehat{\mathbf{u}}_k\\ &+ [(1 - \cos\Theta_k)\mathrm{t}_k\mathrm{w}_k + \sin\Theta_k\mathrm{r}_k]\widehat{\mathbf{f}}_k\\ &+ [(1 - \cos\Theta_k)\mathrm{r}_k\mathrm{w}_k - \sin\Theta_k\mathrm{t}_k]\widehat{\mathbf{v}}_k.\end{aligned} \tag{C7}$$

This relation, together with the two relations obtained from $\widehat{\mathbf{f}}_{k+1} = \mathcal{R}_k\widehat{\mathbf{f}}_k$ and $\widehat{\mathbf{v}}_{k+1} = \mathcal{R}_k\widehat{\mathbf{v}}_k$ can be cast in a matrix product form as

$$\begin{pmatrix}\widehat{\mathbf{f}}_{k+1}\\ \widehat{\mathbf{v}}_{k+1}\\ \widehat{\mathbf{u}}_{k+1}\end{pmatrix} = \mathbf{R}_k\begin{pmatrix}\widehat{\mathbf{f}}_k\\ \widehat{\mathbf{v}}_k\\ \widehat{\mathbf{u}}_k\end{pmatrix}. \tag{C8}$$

The 3×3 matrix $\mathbf{R}_k$ is given by

$$\mathbf{R} = \begin{pmatrix}\cos\Theta + (1 - \cos\Theta)\,t^2 & (1 - \cos\Theta)t\,r + \sin\Theta\,w & (1 - \cos\Theta)t\,w - \sin\Theta\,r\\ (1 - \cos\Theta)t\,r - \sin\Theta\,w & \cos\Theta + (1 - \cos\Theta)\,r^2 & (1 - \cos\Theta)r\,w + \sin\Theta\,t\\ (1 - \cos\Theta)t\,w + \sin\Theta\,r & (1 - \cos\Theta)r\,w - \sin\Theta\,t & \cos\Theta + (1 - \cos\Theta)\,w^2\end{pmatrix}, \tag{C9}$$

where for simplicity we dropped the index k. Setting $k = m - 1$, Eq. (C7) reads

$$\widehat{\mathbf{u}}_m = (\mathbf{R}_{m-1})_{31}\widehat{\mathbf{f}}_{m-1} + (\mathbf{R}_{m-1})_{32}\widehat{\mathbf{v}}_{m-1} + (\mathbf{R}_{m-1})_{33}\widehat{\mathbf{u}}_{m-1}, \tag{C10}$$

a relation that can be iterated further using $\widehat{\mathbf{f}}_{m-1} = \mathcal{R}_{m-2}\widehat{\mathbf{f}}_{m-2}$, $\widehat{\mathbf{v}}_{m-1} = \mathcal{R}_{m-2}\widehat{\mathbf{v}}_{m-2}$, $\widehat{\mathbf{u}}_{m-1} = \mathcal{R}_{m-2}\widehat{\mathbf{u}}_{m-2}$ and similar relations for $m - 2$, $m - 3$.... In this way, one expresses $\widehat{\mathbf{u}}_m$ as a linear combination of $\{\widehat{\mathbf{f}}_0, \widehat{\mathbf{v}}_0, \widehat{\mathbf{u}}_0\}$ with coefficients given as products of rotation matrices (C9). The tangent-tangent correlator (33) then becomes element 33 of the product of these matrices

$$\mathcal{C}_B(m) = \langle\widehat{\mathbf{u}}_0 \cdot \widehat{\mathbf{u}}_m\rangle = \langle\mathbf{R}_{m-1}\dots\mathbf{R}_1\mathbf{R}_0\rangle_{33}. \tag{C11}$$

Next, we develop two approximations for the calculation of $\mathcal{C}_B(m)$. The first one assumes the rotation angle Θ to be infinitesimal. The second one, which is a better approximation, relies on the fact that for DNA, the rotation from one base pair attached triad to the next is dominated by the intrinsic twist component.

1. Infinitesimal rotations

We consider the limit $\Theta \to 0$ and develop $\cos\Theta$ and $\sin\Theta$ in (C9) to lowest order in Θ. Formally, this can also be considered as the continuum limit $a \to 0$, which gives to lowest order [using (C6)]:

$$\begin{aligned}\mathbf{R}_{33} &= 1 - \frac{\Theta^2}{2}(1 - w^2) = 1 - \frac{\Theta^2}{2}(t^2 + r^2)\\ &= 1 - \frac{a^2}{2}(\tau^2 + \rho^2).\end{aligned} \tag{C12}$$

Likewise, $\mathbf{R}_{13} \approx -\mathbf{R}_{31} \approx -a\rho$, $\mathbf{R}_{23} \approx -\mathbf{R}_{32} \approx a\tau$ and similar expressions for the other elements. We consider next the product between two rotation matrices to lowest order in a. For instance, for element 13, we get

$$\begin{aligned}(\mathbf{R}_1\mathbf{R}_0)_{13} &= (\mathbf{R}_1)_{11}(\mathbf{R}_0)_{13} + (\mathbf{R}_1)_{12}(\mathbf{R}_0)_{23} + (\mathbf{R}_1)_{13}(\mathbf{R}_0)_{33} \\ &= -a(\rho_1 + \rho_0) + O(a^2). \end{aligned} \tag{C13}$$

We notice that when calculating this product, we can set $(\mathbf{R}_1)_{11} = 1$ and $(\mathbf{R}_1)_{12} = 0$ as their higher order corrections in a do not contribute to the lowest order in a to the end result in (C13). Analogously, when computing $(\mathbf{R}_1\mathbf{R}_0)_{23}$ we can set $(\mathbf{R}_1)_{21} = 0$ and $(\mathbf{R}_1)_{22} = 1$. Summarizing, if one is interested in the 33 entry of the product of rotation matrices as in (C11) to lowest order in a, it is sufficient to approximate a rotation matrix as

$$\mathbf{R}_n = \begin{pmatrix} 1 & 0 & -a\rho_n \\ 0 & 1 & a\tau_n \\ a\rho_n & -a\tau_n & 1 - \frac{a^2}{2}\left(\tau_n^2 + \rho_n^2\right) \end{pmatrix}. \tag{C14}$$

The product of two such matrices (again to lowest order in a) gives

$$\mathbf{R}_1\mathbf{R}_0 = \begin{pmatrix} 1 & 0 & -a(\rho_1 + \rho_0) \\ 0 & 1 & a(\tau_1 + \tau_0) \\ a(\rho_1 + \rho_0) & -a(\tau_1 + \tau_0) & X_{0,1} \end{pmatrix}, \tag{C15}$$

where we defined

$$\begin{aligned} X_{0,1} &= \left[1 - \frac{a^2}{2}\left(\tau_1^2 + \rho_1^2\right)\right]\left[1 - \frac{a^2}{2}\left(\tau_0^2 + \rho_0^2\right)\right] \\ &\quad - a^2\tau_0\tau_1 - a^2\rho_0\rho_1 \\ &= 1 - \frac{a^2}{2}[(\tau_0 + \tau_1)^2 + (\rho_0 + \rho_1)^2] + O(a^4). \end{aligned} \tag{C16}$$

In conclusion, the product yields again a matrix of the form (C14) with tilt and roll given as the sum of the tilt and roll of the two matrices. This can be generalized to the product of m matrices

$$(\mathbf{R}_{m-1}\dots\mathbf{R}_1\mathbf{R}_0)_{33} = 1 - \frac{a^2}{2}\left[\left(\sum_{k=0}^{m-1}\tau_k\right)^2 + \left(\sum_{k=0}^{m-1}\rho_k\right)^2\right]. \tag{C17}$$

Combining this last result and Eq. (34), we get

$$\frac{1}{l_\mathrm{B}} = \frac{a}{2m}\left\langle\left(\sum_{k=0}^{m-1}\tau_k\right)^2 + \left(\sum_{k=0}^{m-1}\rho_k\right)^2\right\rangle, \tag{C18}$$

which, as done for the torsional persistence length (31), in the limit $N \to \infty$ can be written as

$$\frac{1}{l_\mathrm{B}} = \frac{a}{\pi m}\int_{-\pi/2}^{\pi/2} \frac{\sin^2 my}{\sin^2 y}\,\frac{\langle|\widetilde{\tau}_q|^2 + |\widetilde{\rho}_q|^2\rangle}{N}\,dy, \tag{C19}$$

where as in the main text $y = \pi q/N$.

2. Intrinsic twist dominance

An improved approximation scheme uses the fact that the rotation is dominated by the intrinsic twist component. Indeed, in DNA one has $\omega_0 \gg |\Omega|, |\tau|, |\rho|$, where the difference is typically one order of magnitude. In degrees (note that $a\tau$, $a\rho$, $a\Omega$ are otherwise given in radians), the intrinsic twist angle is $a\omega_0 \approx 34^\circ$, while the other angles are a few degrees. This suggests that one can decompose

$$\mathbf{R}_n = \mathbf{S}\widehat{\mathbf{R}}_n, \tag{C20}$$

as the product of two rotations where $\widehat{\mathbf{R}}_n$ is small and $\mathbf{S}$ a pure twist rotation of magnitude $a\omega_0$. Setting $\mathrm{t} = \mathrm{r} = 0$, $\mathrm{w} = 1$ and $\Theta = a\omega_0$ in (C9), we have

$$\mathbf{S} = \begin{pmatrix} \cos(a\omega_0) & \sin(a\omega_0) & 0 \\ -\sin(a\omega_0) & \cos(a\omega_0) & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{C21}$$

The product of two consecutive rotation matrices is

$$\mathbf{R}_1\mathbf{R}_0 = \mathbf{S}^2(\mathbf{S}^{-1}\widehat{\mathbf{R}}_1\mathbf{S})\widehat{\mathbf{R}}_0 = \mathbf{S}^2\mathbf{R}_1^*\mathbf{R}_0^*, \tag{C22}$$

where we defined

$$\mathbf{R}_n^* \equiv (\mathbf{S}^{-1})^n\widehat{\mathbf{R}}_n\mathbf{S}^n = (\mathbf{S}^{-1})^{n+1}\mathbf{R}_n\mathbf{S}^n. \tag{C23}$$

For the product of m matrices, we get

$$\mathbf{R}_{m-1}\dots\mathbf{R}_1\mathbf{R}_0 = \mathbf{S}^m\mathbf{R}_{m-1}^*\dots\mathbf{R}_1^*\mathbf{R}_0^*. \tag{C24}$$

Taking the thermal average of the 33 component of the two sides of the previous equation, we find

$$\mathcal{C}_\mathrm{B}(m) = \langle\mathbf{R}_{m-1}\dots\mathbf{R}_1\mathbf{R}_0\rangle_{33} = \langle\mathbf{R}_{m-1}^*\dots\mathbf{R}_1^*\mathbf{R}_0^*\rangle_{33}, \tag{C25}$$

where we used $(\mathbf{S}^m)_{3k} = \delta_{3k}$. To calculate the bending persistence length, we will be using the right hand side of (C25). Intrinsic twist dominance implies that in (C9) $\mathrm{w} \approx 1$ and $|\mathrm{t}|, |\mathrm{r}| \ll 1$ and $\Theta \approx a\omega_0$. We can use the approximations

$$\mathrm{w} = \sqrt{1 - \mathrm{t}^2 - \mathrm{r}^2} \approx 1 - \frac{\mathrm{t}^2 + \mathrm{r}^2}{2} = 1 + O(\mathrm{t}^2, \mathrm{r}^2), \tag{C26}$$

and $\Theta = a\omega_0 + O(\mathrm{t}^2, \mathrm{r}^2)$. This implies that (C9) to lowest orders in t and r becomes

$$\mathbf{R} = \begin{pmatrix} \cos(a\omega_0) & \sin(a\omega_0) & (1 - \cos(a\omega_0))t - \sin(a\omega_0)r \\ -\sin(a\omega_0) & \cos(a\omega_0) & (1 - \cos(a\omega_0))r + \sin(a\omega_0)t \\ (1 - \cos(a\omega_0))t + \sin(a\omega_0)r & (1 - \cos(a\omega_0))r - \sin(a\omega_0)t & 1 - (1 - \cos(a\omega_0))(t^2 + r^2) \end{pmatrix}. \tag{C27}$$

Note that taking $a \to 0$, one recovers the infinitesimal form (C14). As in that case, we can ignore terms dependent on τ, ρ (t and r) in the upper 2×2 block as these will not contribute to the bending persistence length to significant order. Next, we calculate $\mathbf{R}_n^*$ using the above form of $\mathbf{R}_n$ (C27) and Eq. (C23). The matrices $\mathbf{S}^n$ and $(\mathbf{S}^{-1})^{n+1}$ have a block-diagonal form as (C21) and correspond to a counterclockwise twist rotation of an angle $na\omega_0$ and a clockwise twist rotation of an angle $(n+1)a\omega_0$, respectively. Equation (C23) gives

$$\mathbf{R}_n^* = \begin{pmatrix} 1 & 0 & -a\rho_n^* \\ 0 & 1 & a\tau_n^* \\ a\rho_n^* & -a\tau_n^* & 1 - \frac{a^2}{2}[(\tau_n^*)^2 + (\rho_n^*)^2] \end{pmatrix}, \tag{C28}$$

where

$$\tau_n^* \equiv \frac{s_{n+1} - s_n}{a\omega_0}\tau_n + \frac{c_{n+1} - c_n}{a\omega_0}\rho_n \tag{C29}$$

$$\rho_n^* \equiv \frac{s_{n+1} - s_n}{a\omega_0}\rho_n - \frac{c_{n+1} - c_n}{a\omega_0}\tau_n, \tag{C30}$$

with

$$c_n \equiv \cos(na\omega_0) \quad s_n \equiv \sin(na\omega_0). \tag{C31}$$

In the limit $a \to 0$, one has $c_{n+1} - c_n \sim O(a^2)$ and $s_{n+1} - s_n \approx a\omega_0$, hence $\tau_n^* \approx \tau_n$ and $\rho_n^* \approx \rho_n$ as expected. The matrix (C28) is formally identical to (C14) with the fields τ and ρ replaced by τ^* and ρ^*. The bending persistence length is then given by the analogous of Eq. (C19):

$$\frac{1}{l_{\mathrm{B}}} = \frac{a}{\pi m}\int_{-\pi/2}^{\pi/2} \frac{\sin^2 my}{\sin^2 y}\,\frac{\langle|\widetilde{\tau}^*{}_q|^2 + |\widetilde{\rho}^*{}_q|^2\rangle}{N}\,dy. \tag{C32}$$

Using (C29) and (C30), the Fourier transforms $\widetilde{\tau}^*{}_q$ and $\widetilde{\rho}^*{}_q$ can be expressed in terms of the original fields. The calculation of the averages in (C32) gives

$$\langle|\widetilde{\tau}^*{}_q|^2 + |\widetilde{\rho}^*{}_q|^2\rangle = \frac{1 - \cos(a\omega_0)}{a^2\omega_0^2}\langle|\widetilde{\tau}_{q+\Delta q}|^2 + |\widetilde{\tau}_{q-\Delta q}|^2 + |\widetilde{\rho}_{q+\Delta q}|^2 + |\widetilde{\rho}_{q-\Delta q}|^2\rangle, \tag{C33}$$

where $\Delta q \equiv Na\omega_0/2\pi$ is the momentum shift associated with the double helix periodicity and originates from the Fourier transforms of c_n and s_n in (C29) and (C30). Combining (C32) and (C33), one obtains the expression of the persistence length (35) given in the main text.

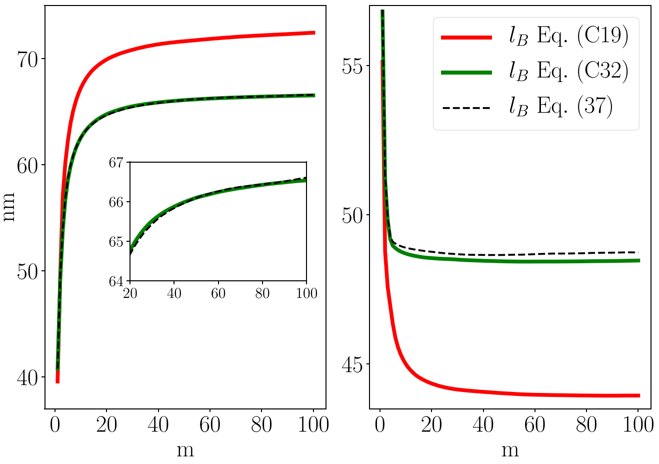

FIG. 8. Monte Carlo simulations with positive (left) and negative (right) off-diagonal couplings. In both cases, couplings between step parameters up to two steps displaced were included. The black lines show the bending persistence length as deduced directly from the tangent-tangent correlation function [Eq. (34)]. Indicated in red is the expression derived for infinitesimal rotations [Eq. (C19)] and in green the improved expression [Eq. (C33)].

TABLE III. Parameters, given in nm, used in the Monte Carlo simulations for the calculation of l_{B} shown in Fig. 8 (X_k indicates the coupling between site n and $n+k$). For the intrinsic twist density and discretization length, $\omega_0 = 1.77$ nm^{-1} and $a = 0.34$ nm were used, respectively.

	Simulation 1			Simulation 2		
	X_0	X_1	X_2	X_0	X_1	X_2
A^t	60	15	5	70	−10	−5
A^r	40	8	4	60	−10	−4
C	80	11	3	100	−20	−5
G	20	2	1	30	−10	−5
A^{tr}	0	−2	0.5	0	0	0
B	0	1	0.5	0	0	0

To compare the quality of these approximations we employed the Monte Carlo method used in Ref. [52] to generate canonical ensembles of triads, distributed according to the free energy (22). In Fig. 8, we compare the direct calculation of the persistence length, as deduced from the tangent-tangent correlation function [Eq. (34)], with the two approximations [Eqs. (C19) and (C33)] for two different sets of model parameters (parameters given in Table III). In both cases, the expression that takes the twist dominance into account [Eq. (C33)], yields excellent agreement with the direct calculation.

APPENDIX D: DETAILS ALL ATOM SIMULATIONS

Using the X3DNA webtool [53], we created an ideal B-DNA duplex structure for various oligomers of 21 and 32 base-pair lengths. All sequences used in this paper are listed in Table IV. The structure was placed in a periodic dodecahedral box with at least 1 nm distance between DNA and box boundary, followed by the addition of water and 150 mM NaCl, resulting in a charge-neutral system (the Debye length is $\lambda_{\mathrm{D}} = 0.8$ nm). Preparation of the system consisted of energy minimization (conjugate gradient with a force threshold of 100 kJ/mol nm) and a 100 ps position restrained molecular dynamics (MD) run, with restraints on the DNA heavy atoms using a force constant of 1000 kJ/mol nm in each direction. We used the parmbsc1 force field [54] to describe the interactions between atoms, in combination with the TIP3P water model [55]. Nonbonded interactions were treated with a cutoff at 1.1 nm, and long range electrostatics were handled by the particle

TABLE IV. Details of the conducted simulations. N is the amount of deformation vectors $\mathbf{\Delta}_n$ considered per snapshot.

Sequence	Simulation time (ns)	N
cgcattgcatacacttggacg	1000	15
cggtaccggctctggtcgccg	1000	15
cgcgatagcgttgtctcaccg	1000	15
cgagttttgaatataagctcg	1000	15
cgggatcaggaaggtggcccg	1000	15
cgttaaagaacatctacgtcg	1000	15
cgatgggcgcggaggcagccg	1000	15
cgtcgagtaacccctaattcg	1000	15
cggcacgggacgaaatcggcg	1000	15
cgactagcatgactgtgcgcg	1000	15
cgttatgtcattataagctcaatgcttatacg	255	27
cgacgtattaccgtacgattggcactatcacg	254	27
cgaagcactgccggggatctgacatccgcgcg	174	27

mesh Ewald method. After equilibration, we performed unrestrained molecular dynamics runs at constant temperature and pressure. The velocity-rescaling thermostat [56] kept the temperature constant at 298 K and the Parrinello-Rahman barostat [57] kept the pressure constant at 1 bar. All molecular dynamics simulations were performed with GROMACS version 2018.6 [58]. Frames were stored every 1 ps. The rotational degrees of freedom of the inter-base-pair parameter—tilt, roll, and twist—were then calculated with the Curves+ algorithm [27]. Figure 9 shows the elements of the stiffness matrix $\widetilde{M}_q$ for the ten different sequences with $N = 15$ and the three sequences with $N = 27$ (corresponding to a 21-mer and 32-mer, respectively), showing some characteristic sample-to-sample variability. The averages of these data are shown in Figs. 5(a) and 5(b).

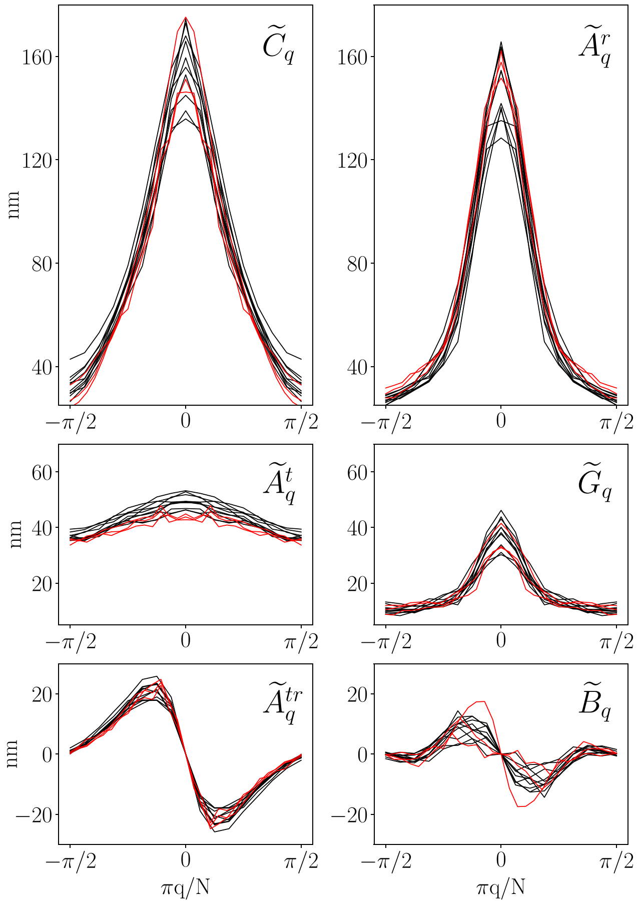

FIG. 9. Entries of the momentum space coupling matrices $\widetilde{M}_q$ for the full spectrum of rescaled momenta for all individual simulations. Results of the 21-mer ($N = 15$) and 32-mer ($N = 27$) simulations are plotted in black and red, respectively.

---

[1] A. Aggarwal, S. Naskar, A. K. Sahoo, S. Mogurampelly, A. Garai, and P. K. Maiti, Curr. Op. Struct. Biol. **64**, 42 (2020).

[2] F. Lankaš, J. Šponer, J. Langowski, and T. E. Cheatham, Biophys J. **85**, 2872 (2003).

[3] F. Lankaš, J. Šponer, P. Hobza, and J. Langowski, J. Mol. Biol. **299**, 695 (2000).

[4] R. Lavery, K. Zakrzewska, D. Beveridge, T. C. Bishop, D. A. Case, T. Cheatham, S. Dixit, B. Jayaram, F. Lankas, C. Laughton, J. H. Maddocks, A. Michon, R. Osman, M. Orozco, A. Perez, T. Singh, N. Spackova, and J. Sponer, Nucl. Acids Res. **38**, 299 (2009).

[5] A. Noy and R. Golestanian, Phys. Rev. Lett. **109**, 228101 (2012).

[6] M. Pasi, K. Zakrzewska, J. H. Maddocks, and R. Lavery, Nucl. Acids Res. **45**, 4269 (2017).

[7] F. Cleri, F. Landuzzi, and R. Blossey, PLoS Comput. Biol. **14**, e1006224 (2018).

[8] V. Velasco-Berrelleza, M. Burman, J. W. Shepherd, M. C. Leake, R. Golestanian, and A. Noy, Phys. Chem. Chem. Phys. **22**, 19254 (2020).

[9] E. Sambriski, D. Schwartz, and J. De Pablo, Biophys. J. **96**, 1675 (2009).

[10] P. D. Dans, A. Zeida, M. R. Machado, and S. Pantano, J. Chem. Theory Comput. **6**, 1711 (2010).

[11] T. E. Ouldridge, A. A. Louis, and J. P. K. Doye, Phys. Rev. Lett. **104**, 178101 (2010).

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

[13] A. Fathizadeh, B. Eslami-Mossallam, and M. R. Ejtehadi, Phys. Rev. E **86**, 051907 (2012).

[14] R. Frederickx, T. in't Veld, and E. Carlon, Phys. Rev. Lett. **112**, 198102 (2014).

[15] Y. A. G. Fosado, D. Michieletto, J. Allan, C. Brackley, O. Henrich, and D. Marenduzzo, Soft Matter **12**, 9458 (2016).

[16] E. Skoruppa, M. Laleman, S. Nomidis, and E. Carlon, J. Chem. Phys. **146**, 214902 (2017).

[17] D. Chakraborty, N. Hori, and D. Thirumalai, J. Chem. Theory Comput. **14**, 3763 (2018).

[18] H. Li and A. Kabakçıoğlu, Phys. Rev. Lett. **121**, 138101 (2018).

[19] E. Skoruppa, S. K. Nomidis, J. F. Marko, and E. Carlon, Phys. Rev. Lett. **121**, 088101 (2018).

[20] O. Henrich, Y. A. G. Fosado, T. Curk, and T. E. Ouldridge, Eur. Phys. J. E **41**, 57 (2018).

[21] M. Caraglio, E. Skoruppa, and E. Carlon, J. Chem. Phys. **150**, 135101 (2019).

[22] P. Nelson, M. Radosavljevic, and S. Bromberg, *Biological Physics: Energy, Information, Life* (W.H. Freeman and Co., New York, 2002).

[23] F. Lankaš, O. Gonzalez, L. Heffler, G. Stoll, M. Moakher, and J. H. Maddocks, Phys. Chem. Chem. Phys. **11**, 10565 (2009).

[24] B. Eslami-Mossallam and M. Ejtehadi, J. Chem. Phys. **134**, 03B623 (2011).

[25] T. Schindler, A. González, R. Boopathi, M. M. Roda, L. Romero-Santacreu, A. Wildes, L. Porcar, A. Martel, N. Theodorakopoulos, S. Cuesta-López *et al.*, Phys. Rev. E **98**, 042417 (2018).

[26] J. Marko and E. Siggia, Macromolecules **27**, 981 (1994).

[27] R. Lavery, M. Moakher, J. Maddocks, D. Petkeviciute, and D. Zakrzewska, Nucl. Acids Res. **37**, 5917 (2009).

[28] S. K. Nomidis, M. Caraglio, M. Laleman, K. Phillips, E. Skoruppa, and E. Carlon, Phys. Rev. E **100**, 022402 (2019).

[29] $\tilde{M}_q$ is Hermitian because $\tilde{\mathbf{\Delta}}_q^\dagger \tilde{M}_q \tilde{\mathbf{\Delta}}_q$ is real for every q.

[30] W. K. Olson, A. A. Gorin, X.-J. Lu, L. M. Hock, and V. B. Zhurkin, Proc. Natl. Acad. Sci. USA **95**, 11163 (1998).

[31] N. Srinivas, T. E. Ouldridge, P. Šulc, J. M. Schaeffer, B. Yurke, A. A. Louis, J. P. Doye, and E. Winfree, Nucl. Acids Res. **41**, 10641 (2013).

[32] T. J. Schmitt, J. B. Rogers, and T. A. Knotts IV, J. Chem. Phys. **138**, 01B613 (2013).

[33] C. Matek, T. E. Ouldridge, J. P. K. Doye, and A. A. Louis, Sci. Rep. **5**, 7655 (2015).

[34] F. Romano and F. Sciortino, Phys. Rev. Lett. **114**, 078104 (2015).

[35] M. C. Engel, D. M. Smith, M. A. Jobst, M. Sajfutdinow, T. Liedl, F. Romano, L. Rovigatti, A. A. Louis, and J. P. Doye, ACS Nano **12**, 6734 (2018).

[36] P. R. Desai, S. Das, and K. C. Neuman, Biophys. J. **118**, 221a (2020).

[37] H. Chhabra, G. Mishra, Y. Cao, D. Prešern, E. Skoruppa, M. Tortora, and J. P. Doye, J. Chem. Theor. Comput. **16**, 7748 (2020).

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

[39] Alternatively, one can construct a global $3N \times 3N$ stiffness matrix and extract the real-space couplings from that analysis. We verified that the results are the same.

[40] Note that the coupling (7) of the toy model can also be expressed as a Fourier series (42) as follows: $\tilde{K}_q = K + 2K' + 2K' \cos(2\pi q/N)$.

[41] We note that the $q = 0$ component is $\mathbf{\Delta}_{q=0} = (\sum_n \tau_n, \sum_n \rho_n, \sum_n \Omega_n)$. The method introduced in Ref. [16] derived asymptotic stiffnesses using a covariance matrix obtained from the sums of τ_n, ρ_n, and Ω_n truncated to an increasing number of terms. Hence, the results reported in Ref. [16] report the $q = 0$ component of the stiffness matrix.

[42] J. Lipfert, G. M. Skinner, J. M. Keegstra, T. Hensgens, T. Jager, D. Dulin, M. Köber, Z. Yu, S. P. Donkers, F.-C. Chou, R. Das, and N. H. Dekker, Proc. Natl. Acad. Sci. USA **111**, 15408 (2014).

[43] E. Trifonov, R. Tan, and S. Harvey, in *Structure and Expression: Proceedings of the Fifth Conversation in the Discipline Biomolecular Stereodynamics*, edited by M. Sarma and R. Sarma (Adenine Press, Schenectady, NY, 1988).

[44] J. S. Mitchell, J. Glowacki, A. E. Grandchamp, R. S. Manning, and J. H. Maddocks, J. Chem. Theory Comput. **13**, 1539 (2017).

[45] S. K. Nomidis, F. Kriegel, W. Vanderlinden, J. Lipfert, and E. Carlon, Phys. Rev. Lett. **118**, 217801 (2017).

[46] S. Kim, E. Broströmer, D. Xing, J. Jin, S. Chong, H. Ge, S. Wang, C. Gu, L. Yang, Y. Q. Gao *et al.*, Science **339**, 816 (2013).

[47] G. Rosenblum, N. Elad, H. Rozenberg, F. Wiggers, and H. Hofmann, bioRxiv (2020), doi: 10.1101/2020.07.04.187450.

[48] D. M. Chenoweth and P. B. Dervan, Proc. Nat. Acad. Sci. USA **106**, 13175 (2009).

[49] T. Drsata, M. Zgarbova, N. Spackova, P. Jurecka, J. Sponer, and F. Lankas, J. Phys. Chem. Letters **5**, 3831 (2014).

[50] J. Singh and P. K. Purohit, J. Phys. Chem. B **123**, 21 (2018).

[51] W. G. Noid, J.-W. Chu, G. S. Ayton, V. Krishna, S. Izvekov, G. A. Voth, A. Das, and H. C. Andersen, J. Chem. Phys. **128**, 244114 (2008).

[52] S. K. Nomidis, E. Skoruppa, E. Carlon, and J. F. Marko, Phys. Rev. E **99**, 032414 (2019).

[53] S. Li, W. Olson, and X.-J. Lu, Nucl. Acids Res. **47**, W26 (2019).

[54] I. Ivani, P. Dans, A. Noy, A. Pérez, I. Faustino, A. Hospital, J. Walther, P. Andrio, R. Goñi, A. Balaceanu, G. Portella, F. Battistini, J. Gelpí, C. González, M. Vendruscolo, C. Laughton, S. Harris, D. Case, and M. Orozco, Nat. Methods **13**, 55 (2016).

[55] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).

[56] G. Bussi, D. Donadio, and M. Parrinello, J. Chem. Phys. **126**, 014101 (2007).

[57] M. Parrinello and A. Rahman, J. Appl. Phys. **52**, 7182 (1981).

[58] M. Abrahams, T. Murtola, R. Schulz, S. Páll, J. Smith, B. Hess, and E. Lindahl, SoftwareX **1-2**, 19 (2015).