

# Density–Nematic Coupling in Isotropic Solution of DNA: Multiscale Model

Daniel Svenšek,\* Jaka Sočan, and Matej Praprotnik

Monte Carlo simulations of isotropic solutions of double-stranded DNA (deoxyribonucleic acid) are performed using the well-established oxDNA model. By comparing the fluctuation amplitudes with theoretical predictions, the parameters of a generic macroscopic model of an isotropic linear polymer solution/melt are determined. A multiscale continuum field model is thus obtained, corresponding to the full specificity of the isotropic phase of double-stranded DNA in the usual B-form as perceived at the macroscopic level. Present research is particularly focused on the coupling between spatial concentration/density variations of the polymer and the emerging nematic orientation order of the chains. This rather unfamiliar, only recently described phenomenon, inherent to linear polymers, is outlined and interpreted. Quantitative predictions are provided for the degree of nematic order induced by concentration gradients in isotropic solutions of double-stranded DNA.

where  $\rho_s(\mathbf{r})$  is the surface number density of polymer chains perforating the plane perpendicular to  $\mathbf{n}$  and  $\rho^+(\mathbf{r})$  and  $\rho^-(\mathbf{r})$  are volume densities of the beginnings and endings of the chains, which fill the voids between the chains created by splay, **Figure 1**, and thus weaken the constraint Equation (1) on  $\rho_s$  and  $\mathbf{n}$ .

Nevertheless, hairpin turns of chains are a characteristic feature of nematic order, and despite the awareness of this fact and the fact that Equation (1) seems to be clearly vectorial in nature, it was formally shown only a decade ago that Equation (1) generalizes to a rigorous continuity equation for the polar order vector  $\mathbf{a}(\mathbf{r})$  instead of the apolar nematic director,<sup>[12,13]</sup>

$$\nabla \cdot (\rho l_0 \mathbf{a}) = \rho^+ - \rho^- \quad (2)$$

## 1. Introduction

In solutions/melts of linear polymers, gradients of orientational order of polymer chains are coupled to gradients of their concentration/density. This macroscopic consequence of the simple existence of microscopic chains is known as splay–density coupling in nematic polymers, where the splay distortion of the director leads to a density gradient and vice versa.

Traditionally, this phenomenon has been described by the Meyer-de Gennes continuity constraint<sup>[1–10]</sup> for the nematic director field  $\mathbf{n}(\mathbf{r})$  – direction of the polymer chain, valid in the absence of hairpins (sharp, ideally point-like 180° backfolds of the chain<sup>[11]</sup>),

$$\nabla \cdot (\rho_s \mathbf{n}) = \rho^+ - \rho^- \quad (1)$$

D. Svenšek, J. Sočan, M. Praprotnik  
 Laboratory for Molecular Modeling  
 National Institute of Chemistry  
 Ljubljana SI-1001, Slovenia  
 E-mail: [daniel.svensek@fmf.uni-lj.si](mailto:daniel.svensek@fmf.uni-lj.si)

D. Svenšek, M. Praprotnik  
 Department of Physics  
 Faculty of Mathematics and Physics  
 University of Ljubljana  
 Ljubljana SI-1000, Slovenia

© 2024 The Author(s). Macromolecular Rapid Communications published by Wiley-VCH GmbH. This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

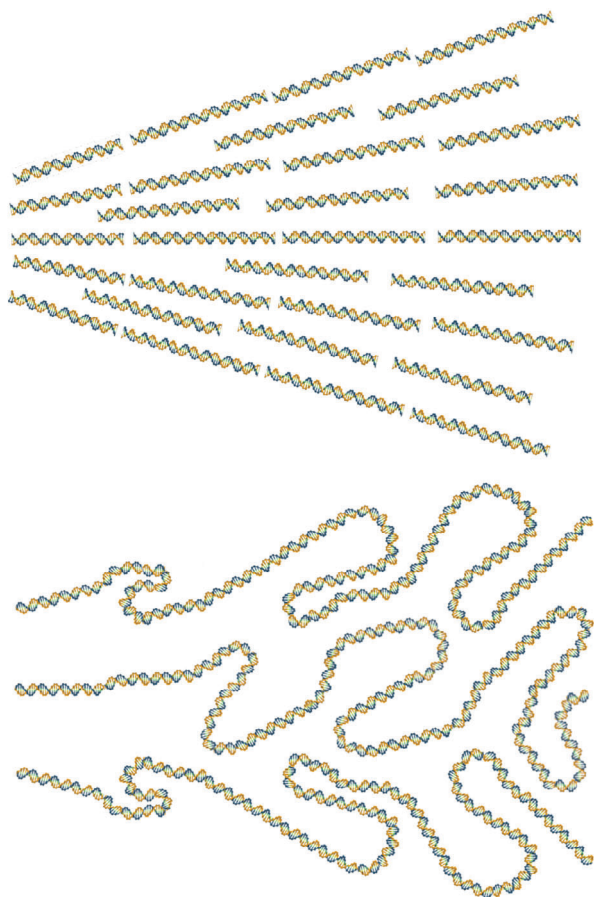
DOI: 10.1002/marc.202400382

where the polymer length per unit volume  $\rho l_0$  is the key quantity. That is,  $l_0$  is the length of arbitrarily (but of course appropriately) chosen chain segments and  $\rho$  is their volume number density, so that the product, the volume density of the polymer length, is a physical fact independent of the choice of the segments. By definition, the polar order vector  $\mathbf{a}$  describes collective (mesoscopic) polar orientational order of the chains, which can be defined if the chains themselves are polar (directed), i.e., composed of polar monomers.

It was also shown how Equation (1) can be formally adapted so that it can be applied to a nonpolar, uniaxial nematic case with an arbitrary density of hairpins. This is where the so-called recovered polar order and new effective chain ends with double “strength” corresponding to hairpins come into play (see refs. [13, 14] for the full story). In the limit of large concentration of hairpins,<sup>[15]</sup> however, the continuity equation for the recovered polar order, although still valid, reduces to a trivial identity and therefore loses its meaning.

Nematic orientational order, which is quadrupolar rather than polar, is described by the symmetric traceless nematic order tensor (the quadrupole moment of the orientational distribution function)  $Q_{ij}(\mathbf{r}) = \frac{3}{2} \left( \langle t_i t_j \rangle - \frac{1}{3} \delta_{ij} \right)$ , where, in the case of linear polymers, the averaging  $\langle \rangle$  is over unit vectors  $\mathbf{t}$  indicating the directions of the chain segments;  $\delta_{ij}$  is the Kronecker delta. The director  $\mathbf{n}$  appearing in Equation (1) is then the principal axis of  $Q_{ij}$ . A rigorous conservation law, analogous to Equation (2) for polar order, was derived for quadrupolar (nematic) order,<sup>[16]</sup>

$$\partial_j \left[ \rho l_0 \left( Q_{ij} + \frac{1}{2} \delta_{ij} \right) \right] = \frac{3}{2} g_i + \frac{3}{2} \rho l_0 k_i \quad (3)$$



**Figure 1.** Schematics of splay deformation in a nematic phase. The density of long chains inevitably decreases as the chains spread. Ensembles of shorter chains (top) contain more chain ends that can fill the voids more easily, thus mitigating the decrease in density. Backfolds (bottom) also fill the voids, but reduce orientational ordering unless they are pointlike U turns (hairpins), which are favored in the nematic phase. In this case, they act as chain ends and can fill the voids similar to the top picture.

where the volume density  $g(\mathbf{r})$  of chain end directions  $\mathbf{t}$ , defined as pointing inward, and the volume density ( $\rho l_0 \mathbf{k}(\mathbf{r})$ ) of deflections of consecutive segments play the role of the sources in this continuity equation;  $\mathbf{k}(\mathbf{r})$  is the local average chain curvature vector. The average chain curvature vector source reflects the effect of general chain folds (including hairpins as their special case), which can fill the voids created by splay in a similar way as the chain ends do, Figure 1 (bottom). The stiffer and longer the chains, the more expensive are the sources and the stronger is the coupling Equation (3) between gradients of  $Q_{ij}$  and  $\rho$ . Analogous to the polar case Equation (2), the tensorial conservation law Equation (3) is an exact macroscopic implication of the simple existence of the microscopic polymer chains.<sup>[17]</sup>

Somewhat unexpectedly, however, Equation (3) is relevant not only for nematic linear polymers, but also for isotropic linear polymers, whether or not they exhibit an isotropic–nematic transition. That is, it applies to any fluid linear polymer (i.e., solution or melt), e.g., synthetic polymers such as polyethylene, polyvinyls, polyamides, polyesters, polystyrene, polycarbonates,

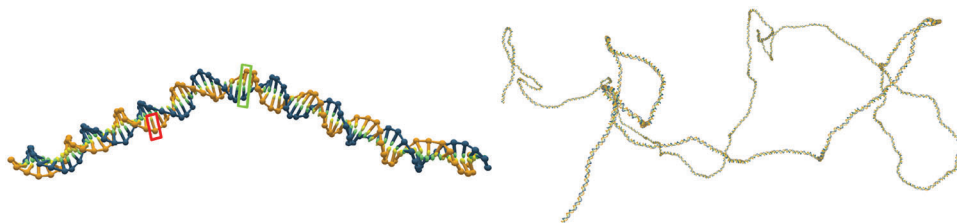
etc., and also to a solution of deoxyribonucleic/ribonucleic acid (DNA/RNA) and other linear biopolymers. Why this relevance? Unlike polar order fluctuations, variations of nematic order  $\delta Q_{ij}$  are generally coupled in the lowest order to variations of polymer density/concentration  $\delta\rho$  even in an orientationally disordered, isotropic phase with  $\rho = \rho_0$  and  $Q_{ij} = 0$  in equilibrium, as follows from linearization of Equation (3):

$$\rho_0 l_0 \partial_j \delta Q_{ij} + \frac{1}{2} l_0 \partial_i \delta \rho = \frac{3}{2} \delta g_i + \frac{3}{2} \rho_0 l_0 \delta k_i \quad (4)$$

This applies to thermal fluctuations as well as to arbitrary, e.g., externally imposed variations  $\delta\rho(\mathbf{r})$  and  $\delta Q_{ij}(\mathbf{r})$ . Of course, there must exist a microscopic object that can be oriented in the first place (in a fluid of spherical particles, for example, this is not possible), but in polymers such an object is always the polymer chain itself (i.e., even if it consisted of spherical monomers). Therefore, for example, an externally imposed concentration gradient of DNA (or other linear polymer) will induce nematic ordering via the coupling Equation (4) and thereby also optical anisotropy (birefringence) in the otherwise isotropic solution, as has already been shown for the case of a generic worm-like polymer melt.<sup>[16]</sup>

Such osmotic-stress-induced birefringence is similar to shear flow-induced birefringence in fluid polymers and to the stress-optic law in elastic solid dielectric materials (a direct coupling between strain and dielectric tensors). However, the key difference is that osmotic-stress-induced birefringence occurs already in a static liquid where no strain or strain rate tensor exists to couple with the dielectric tensor. Instead, the coupling, which arises from the microscopic polymer chain connectivity, occurs via the concentration gradient as expressed by the tensorial continuity equation Equation (3).

To the best of our knowledge, birefringence induced by polymer concentration gradients has not yet been evidenced in the literature, but the basic arguments presented above suggest that it should be a very persistent, geometrically enforced phenomenon. To theoretically describe this interesting macroscopic mechanism and other phenomena at the macroscopic level, a macroscopic model of the specific polymer is required, which, in order to achieve specificity, must be based on the microscopic information of the concrete system. Hierarchical multiscale modelling has so far proven to efficiently characterize phenomena, which need to be addressed at various levels of complexity,<sup>[18]</sup> such as surface catalyst optimization,<sup>[19]</sup> composite material architecture determination<sup>[20]</sup> and prediction of multi-principal alloy behavior under stress.<sup>[21]</sup> Here, we employ a multiscale solution that involves a theoretical description of our phenomena at the macroscopic level in conjunction with a micro- to mesoscopic description of our polymer of interest. We set up a minimal macroscopic continuum model for a generic isotropic linear polymer solution or melt and determine its parameters corresponding specifically to the isotropic solution of double-stranded (ds)DNA in B-form at various chain lengths and concentrations, as simulated with the well-established oxDNA model.<sup>[22]</sup> The B-form is the most common isoform of dsDNA, where the planes of the nucleic acid bases are nearly perpendicular to the helical axis. It is the predominant form in biological systems and in vitro aqueous solutions at (quasi) physiological conditions.



**Figure 2.** A representation of dsDNA chains of lengths 80 (left) and 5000 (right) base pairs, as modeled by oxDNA. In the oxDNA coarse-grained force field, each nucleotide (denoted by the red shape) is represented as a rigid body interacting with its surroundings through pairwise interactions, i.e., backbone connectivity and repulsion, as well as nucleobase stacking, repulsion and hydrogen-bonding. In present work, we treat two hydrogen-bound nucleotides – the base pair (denoted by the green shape) – as a dsDNA monomer.

## 2. Multiscale Model

Following,<sup>[16]</sup> we use a minimal macroscopic free-energy density model of an isotropic linear polymer solution or melt that is capable of describing variations of polymer density, nematic orientational order of chains, and coupling between the two:

$$f = \frac{1}{2}B\left(\frac{\delta\rho}{\rho_0}\right)^2 + \frac{1}{2}A(\delta Q_{ij})^2 + \frac{1}{2}L(\partial_k\delta Q_{ij})^2 + \frac{1}{2}G\left(\frac{2}{3}\rho_0 l_0\right)^2 \left[ \partial_j\delta Q_{ij} + \frac{1}{2}\partial_i\left(\frac{\delta\rho}{\rho_0}\right) \right]^2 \quad (5)$$

where the variables are the variation of the relative number density of the segments  $\delta\tilde{\rho} = \delta\rho/\rho_0$ ,  $\rho_0$  is its equilibrium value, and the variation of the nematic order tensor  $\delta Q_{ij}$ , which is zero in equilibrium. The first term describes the cost of density or concentration variations, with  $B$  the bulk modulus (in the case of a melt) or the concentration susceptibility (in the case of a solution). The second term captures the cost of nonzero nematic order, which is disfavored in the isotropic system,  $A$  is the “stiffness” of the nematic order, i.e., the nematic order modulus. The third term (nematic elastic term) penalizes the gradients of the nematic order. In contrast to the nematic phase, where there are several elastic terms, three by default corresponding to splay, twist and bend distortion modes of the nematic director, there is only one elastic term in the isotropic system. This is due to the fact that there is no preferred direction in equilibrium and thus no distinction between different distortion modes of nematic order is possible. The last term describes the density–nematic coupling. It is enforced by a quadratic potential penalizing the sources of the tensorial conservation law, i.e., the right-hand side of Equation (4) expressed by its left-hand side containing the variables;  $G\left(\frac{2}{3}\rho_0 l_0\right)^2 \equiv \tilde{G}$  is the strength of the coupling.

Once such a model is established, it is useful in many ways. It provides amplitudes of thermal fluctuations of polymer density and nematic order, the correlation between these two fluctuations, it describes the effects of externally induced density variations or induced nematic order, it can also be supplemented by contributions from static external fields.

The parameters of the model Equation (5),  $B$ ,  $A$ ,  $L$ ,  $\tilde{G}$ , are unknown and must be determined. A well-established procedure for this is to compare the amplitudes of long-wave fluctuations predicted by Equation (5) with the same fluctuations extracted from the actual experimental system. Traditionally, this is done with different types of scattering experiments (X-ray, neutron,

and light scattering).<sup>[23–27]</sup> Nowadays, a complementary and very convenient in silico multiscale strategy is also possible, i.e., performing microscopic simulations and calculating the fluctuation amplitudes directly from numerical data.<sup>[13,16,28–35]</sup> Here, we will use conformational ensembles of dsDNA produced by oxDNA model<sup>[22]</sup> Monte Carlo (MC) simulations, **Figure 2**.

To obtain the macroscopic fluctuation amplitudes, Equation (5) is expressed in Fourier space,<sup>[36]</sup>

$$f(\mathbf{q}) = \frac{1}{2}B|\delta\tilde{\rho}(\mathbf{q})|^2 + \frac{1}{2}(A + Lq^2)|\delta Q_{ij}(\mathbf{q})|^2 + \frac{1}{2}\tilde{G}|q_j\delta Q_{ij}(\mathbf{q}) + \frac{1}{2}q_i\delta\tilde{\rho}(\mathbf{q})|^2 \quad (6)$$

where

$$\delta\tilde{\rho}(\mathbf{q}) = \int d^3r \delta\tilde{\rho}(\mathbf{r})e^{-iq\cdot\mathbf{r}}, \quad \delta Q_{ij}(\mathbf{q}) = \int d^3r \delta Q_{ij}(\mathbf{r})e^{-iq\cdot\mathbf{r}} \quad (7)$$

Thus, the free energy  $F = \int d^3r f$  is now a sum over the Fourier modes  $F = (1/V) \sum_{\mathbf{q}} f(\mathbf{q})$ , where  $V$  is the volume of the system. By equipartition, the energy corresponding to a stand-alone quadratic contribution  $f_i(\mathbf{q})$  of the energy  $f(\mathbf{q})$  is  $\langle f_i(\mathbf{q}) \rangle / V = k_B T/2$ , with  $k_B$  the Boltzmann constant and  $T$  the temperature. To determine the fluctuation amplitudes, the quadratic form Equation (6) is diagonalized (see [37] Supporting Information of ref. [16]) and expressed by terms purely quadratic in proper linear combinations of the variables, from which the fluctuation amplitudes follow by equipartition. Since the system is isotropic, we can assume without loss of generality  $\mathbf{q} = q\hat{e}_z$  for the fluctuation wave vector, where  $z$  is an arbitrarily chosen direction defining the  $z$  axis of the coordinate system (the results are the same for any choice of  $z$  axis). Axes  $x$  and  $y$  are then arbitrarily chosen in the plane perpendicular to  $z$  and all results at a given  $q$  are invariant to rotations of the tensors in the  $xy$  plane, in particular  $\langle |\delta Q_{xz}|^2 \rangle = \langle |\delta Q_{yz}|^2 \rangle$  and  $\langle |\delta Q_{xx} - \delta Q_{yy}|^2 \rangle = 4\langle |\delta Q_{xy}|^2 \rangle$ . The fluctuation amplitudes are<sup>[16]</sup>

$$\frac{1}{N_0} \langle |\delta Q_{xy}|^2 \rangle = \frac{k_B T}{2} \frac{1}{\rho_0} \frac{1}{A + Lq^2} \quad (8)$$

$$\frac{1}{N_0} \langle |\delta Q_{\{xz,yz\}}|^2 \rangle = \frac{k_B T}{2} \frac{1}{\rho_0} \frac{1}{A + (L + \frac{1}{2}\tilde{G})q^2} \quad (9)$$

$$\frac{1}{N_0} \langle |\delta Q_{zz}|^2 \rangle = \frac{k_B T}{2} \frac{4}{\rho_0} \left[ 3A + \left( 3L + \frac{8\tilde{G}B}{4B + \tilde{G}q^2} \right) q^2 \right]^{-1} \quad (10)$$

$$\frac{1}{N_0} \langle |\delta\tilde{\rho}|^2 \rangle = \frac{k_B T}{2} \frac{8}{\rho_0} \left[ 4B + \frac{3\tilde{G}(A + Lq^2)q^2}{3A + (3L + 2\tilde{G})q^2} \right]^{-1} \quad (11)$$

$$\begin{aligned} & \frac{1}{2N_0} \langle \delta\tilde{\rho}^* \delta Q_{zz} + \delta\tilde{\rho} \delta Q_{zz}^* \rangle \\ &= -\frac{k_B T}{2} \frac{1}{\rho_0} \frac{8\tilde{G}q^2}{12AB + [12BL + (3A + 8B)\tilde{G}]q^2 + 3\tilde{G}Lq^4} \end{aligned} \quad (12)$$

where  $N_0 \equiv \rho_0 V$  is the total number of segments in the system. Dividing the fluctuation amplitudes, Equations (8)–(12), by  $N_0$  is convenient, as in this way they become intensive quantities that do not trivially depend on  $N_0$ . Equation (12) gives the correlation of  $\delta\tilde{\rho}(q)$  and  $\delta Q_{zz}(q)$ , i.e., only the component  $\delta Q_{zz}$  is directly coupled to density. The fluctuation  $\langle |\delta Q_{xy}|^2 \rangle$  is the only one that does not depend on the coupling strength  $\tilde{G}$ , i.e., the difference between  $\langle |\delta Q_{xy}|^2 \rangle$  and  $\langle |\delta Q_{xz}|^2 \rangle = \langle |\delta Q_{yz}|^2 \rangle$  is a clear signature of the density–nematic coupling in the isotropic phase! Note that  $\delta Q_{zz} = -(\delta Q_{xx} + \delta Q_{yy})$  because of the tracelessness of  $Q_{ij}$ .

In Equations (8)–(12), all amplitudes  $\delta Q_{xy}$ ,  $\delta Q_{xz}$ ,  $\delta Q_{yz}$ ,  $\delta Q_{zz}$ , and  $\delta\tilde{\rho}$  of originally dimensionless quantities are Fourier components Equation (7) and therefore have dimensions of  $V$ . In microscopic simulations, on the other hand, the Fourier transform is not performed by volume integration of continuous fields, but by summation over discrete polymer segments, so that the transformed quantities remain dimensionless. To allow direct comparison of the fluctuation amplitudes, we multiply the Equations (8)–(12) by  $\rho_0^2$  so that they express the same dimensionless quantities  $\langle |\rho_0 \delta Q_{xy}|^2 \rangle / N_0, \dots$ , and  $\langle |\rho_0 \delta\tilde{\rho}|^2 \rangle / N_0 = \langle |\delta\rho|^2 \rangle / N_0$ . Moreover, from now on we also use dimensionless units by expressing energy relative to  $k_B T$  and length relative to  $l_0$ . Specifically, the unit of the parameters  $B$ ,  $A$  is  $k_B T / l_0^3$ , and the unit of the parameters  $L$ ,  $\tilde{G}$  is  $k_B T / l_0$ .

With all that, the form of the macroscopic fluctuation amplitudes Equations (8)–(12), which can be compared directly with values obtained from the oxDNA simulations, is

$$\frac{1}{N_0} \langle |\rho_0 \delta Q_{xy}|^2 \rangle = \frac{1}{2} \rho_0 \frac{1}{A + Lq^2} \quad (13)$$

$$\frac{1}{N_0} \langle |\rho_0 \delta Q_{\{xz,yz\}}|^2 \rangle = \frac{1}{2} \rho_0 \frac{1}{A + (L + \frac{1}{2}\tilde{G})q^2} \quad (14)$$

$$\frac{1}{N_0} \langle |\rho_0 \delta Q_{zz}|^2 \rangle = 2\rho_0 \left[ 3A + \left( 3L + \frac{8\tilde{G}B}{4B + \tilde{G}q^2} \right) q^2 \right]^{-1} \quad (15)$$

$$\frac{1}{N_0} \langle |\delta\rho|^2 \rangle = 4\rho_0 \left[ 4B + \frac{3\tilde{G}(A + Lq^2)q^2}{3A + (3L + 2\tilde{G})q^2} \right]^{-1} \quad (16)$$

$$\begin{aligned} & \frac{1}{2N_0} \langle \delta\rho^* \rho_0 \delta Q_{zz} + \delta\rho \rho_0 \delta Q_{zz}^* \rangle \\ &= -4\rho_0 \frac{\tilde{G}q^2}{12AB + [12BL + (3A + 8B)\tilde{G}]q^2 + 3\tilde{G}Lq^4} \end{aligned} \quad (17)$$

### 3. Comparison of Fluctuations, Extracted Model Parameters

The full comparison of the macroscopic fluctuation amplitudes, Equations (13)–(17), relevant in the long-wavelength limit, with the same long-wavelength fluctuations obtained from the sim-

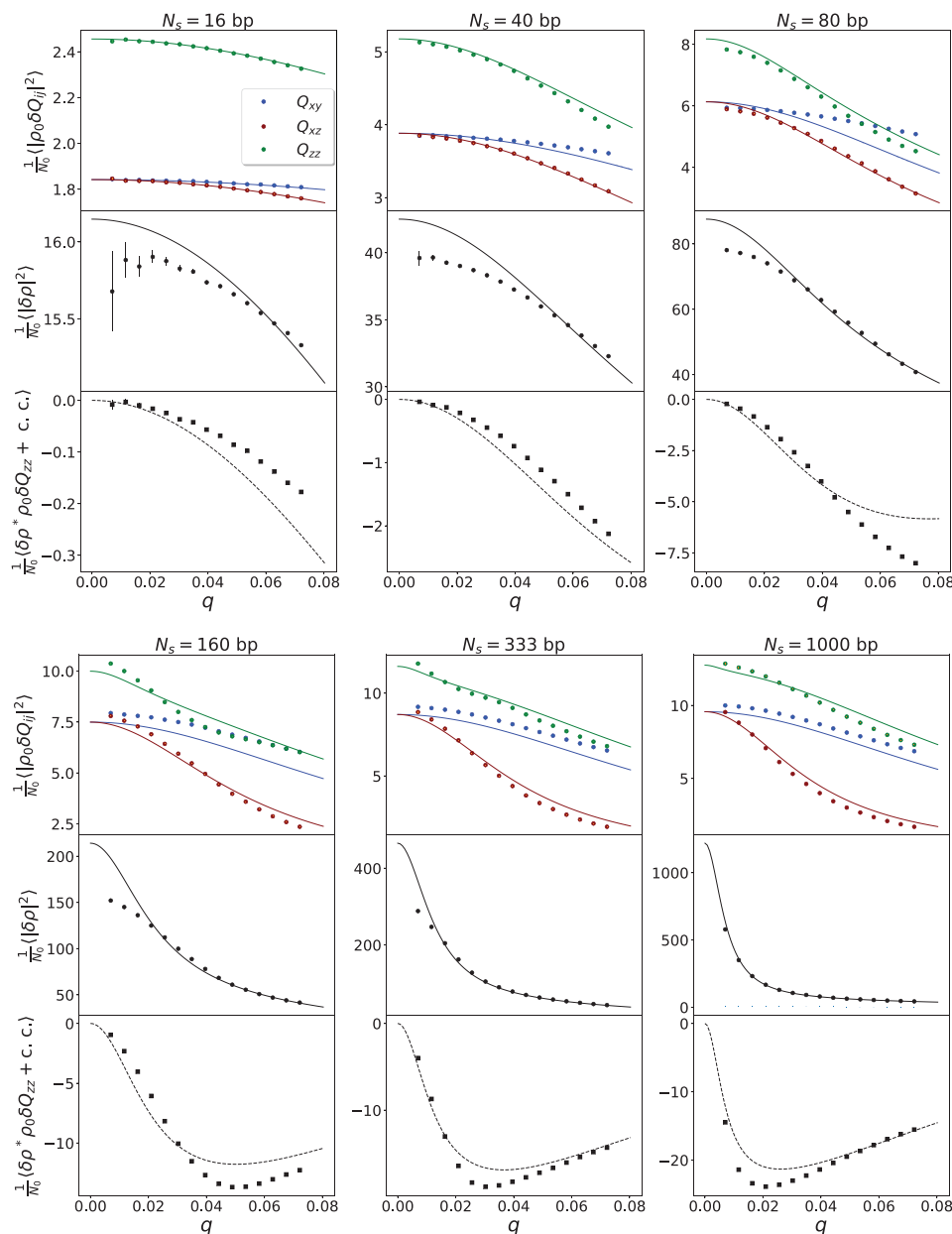
ulations is shown in **Figure 3** for all six chain lengths and at monomer density  $\rho_0 = 5\rho_{00}$  as an example. For a given chain length and monomer density, the fluctuations Equations (13)–(16) are fitted simultaneously to determine the model parameters  $B$ ,  $A$ ,  $L$ ,  $\tilde{G}$  for that chain length and monomer density. The cross-correlation Equation (17) is not fitted but only plotted with the determined parameters. For other simulated densities  $\rho_{00}$ ,  $2\rho_{00}$ ,  $10\rho_{00}$ , and  $20\rho_{00}$ , the  $q$ -dependences of all fluctuations and their agreements/discrepancies with theoretical predictions are very similar. The statistical errors are insignificant everywhere, except for the density fluctuation at the lowest  $q$  points.

Despite the fact that the system is isotropic ( $\langle \mathbf{Q} \rangle = 0$ ) and far from any pretransitional orientational effects, the orientation and density fluctuations indeed reveal the fundamental implications of its linear chain structure. The difference between the  $Q_{xy}$  and  $Q_{xz}$  fluctuations is a direct consequence of the density–nematic coupling and serves as a first indicative measure of the magnitude of the coupling strength  $\tilde{G}$  in Equation (14). Already in **Figure 3**, we can immediately see that it increases systematically with the length of the chains, as expected. According to the discussion after Equation (3) longer chains have fewer ends and consequently the corresponding source  $\mathbf{g}$  is more expensive, i.e., its fluctuation  $\delta\mathbf{g}$  in Equation (4) is smaller and thus the fluctuations of  $\rho$  and  $\mathbf{Q}$  are more strongly coupled.

The  $q$ -dependence of the density fluctuation in **Figure 3** is also due to this coupling – for  $\tilde{G} = 0$ ,  $|\delta\rho|^2$  would be independent of  $q$  in this low- $q$  region. We see that it is again much stronger for longer chains. The ultimate manifestation of the density–nematic coupling is, of course, the nonzero cross-correlation of  $\delta\rho$  and  $\delta Q_{zz}$  shown in the lower diagrams of **Figure 3**. Again, it increases systematically with the length of the chains.

Before turning to the details, let us put the results presented in **Figure 3** in an overall context. First, they clearly show the presence of the density–nematic coupling. Second, they also show some deviations from the macroscopic model Equation (5) that incorporates this coupling in the simplest manner. We should not forget that the continuity equation Equation (3) is exact and not an approximation. Therefore, it will inevitably manifest itself in some form in the variations of density and orientational order. It is just that the energy cost of the sources of this continuity equation cannot be fully described by the universal quadratic potential of the model Equation (5), and hence the observed discrepancies in **Figure 3**.

Our main result is the dependence of the parameters of the model Equation (5) on the polymer chain length  $N_s$  and the density of the monomers  $\rho_0$ , shown in **Figure 4** and in **Table 1**. The expression Equation (5) is a free-energy density and therefore its parameters are proportional to the monomer density  $\rho_0$  in first approximation in the absence of other effects. No peculiarities are expected for the compressibility modulus  $B$  and nematic order stiffness  $A$ , while this is not so clear in advance for the other two terms with gradients. The plot of the ratios  $A/\rho_0$ ,  $L/\rho_0$ ,  $\tilde{G}/\rho_0$ , and  $B/\rho_0$  in **Figure 4** shows that for densities  $\rho_{00}$  to  $20\rho_{00}$  this proportionality is practically exact for all parameters, indicating that the system is essentially dilute (based on crystallographic data,<sup>38</sup> at  $\rho_0 = 20\rho_{00}$  the solute occupies only about 0.012% of the simulation box volume). There are only minor systematic deviations within the error bars for  $L$  and  $\tilde{G}$ . However, noticeable

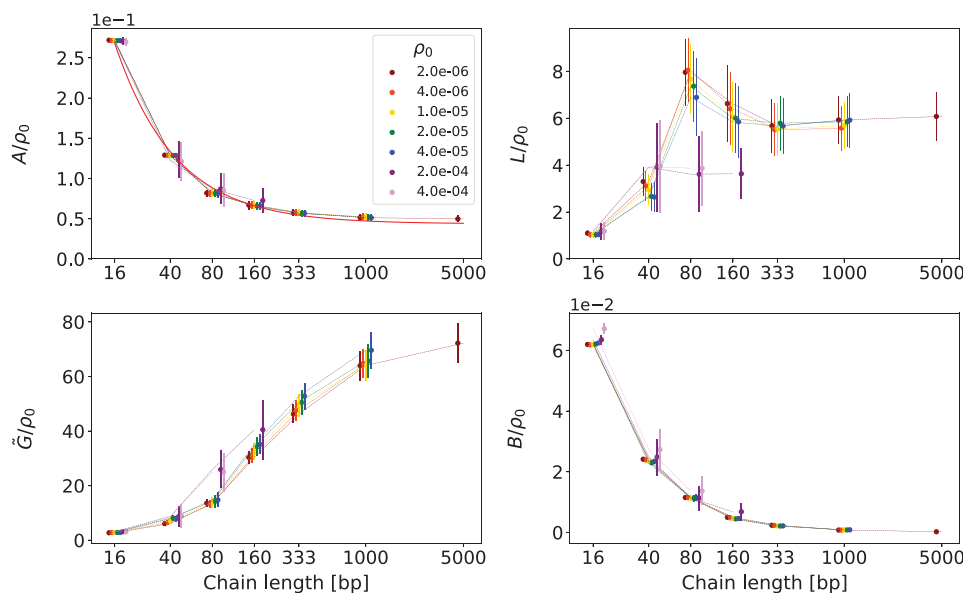


**Figure 3.** Fluctuation amplitudes and fits to Equations (13)–(16) for different chain lengths  $N_s$  and base pair density  $\rho_0 = 5\rho_{00}$  ( $\approx 250000$  base pairs). The cross-correlation curves (dashed) are direct plots of Equation (17) (no fitting). Measured standard deviations are shown by error bars that are too small to be visible except for density fluctuations and cross-correlations of short chains. The lowest  $q > 0$  bin is omitted everywhere because of large statistical errors for short chains.

discrepancies, particularly in  $L/\rho_0$  and  $\tilde{G}/\rho_0$ , appear for the higher densities of  $100\rho_{00}$  and  $200\rho_{00}$ .

The decreasing dependence of the compressibility modulus  $B$  on the chain length, Figure 4, is in agreement with the known theoretical result for an ideal polymer chain (entropic compressibility modulus)<sup>[39], pp. 19–20].<sup>[40]</sup> Namely, in the limiting case of weak confinement, i.e., when the average gyration radius (rms radius)  $r_g$  of the chain is much smaller than the confining box of volume  $V$ , the pressure exerted by the ideal chain is simply  $p \approx k_B T/V$  and the compressibility</sup>

modulus is  $B = (1/\rho)\partial p/\partial \rho = -V\partial p/\partial V \approx k_B T/V \approx p$ . That is, in this limit the ideal chain behaves like a single ideal gas particle. For  $N_c$  independent ideal chains, this means  $B \approx p \approx N_c k_B T/V = \rho_0 k_B T/N_s$ . Thus, in our dimensionless units the ratio  $B/\rho_0$ , which is shown in Figure 4, would ideally be  $B/\rho_0 \approx 1/N_s$ , i.e.,  $\log(B/\rho_0) \approx -\log N_s$ . The log–log plot in Figure 5 indeed shows a quasi-linear dependence with a slope close to  $-1$ . For longer chains, the line rises slightly, which may indicate that  $r_g$  of these chains is no longer very small compared to the size of the box, and their segments begin to make individual



**Figure 4.** The central result of the simulation: the dependence of the parameters of the model Equation (5) on the polymer chain length  $N_s$  (note the logarithmic scale) and the density of the monomers  $\rho_0$ . The proportionality of the nematic order stiffness  $A$  and compressibility  $B$  to  $\rho_0$  is exact. For better visibility, the points corresponding to different densities are slightly shifted horizontally. A curve  $\approx 3.60 (1/N_s + 1/82.5)$  is fitted to the  $A/\rho_0$  data.

contributions to the pressure. For the two higher densities,  $B/\rho_0$  is slightly higher and in particular the lines in Figure 5 rise more strongly, which points to the effect of chain's reduced free volume. This volume, rather than the full volume of the box, is now to be compared to the volume  $4\pi r_g^3/3$  occupied by a chain.

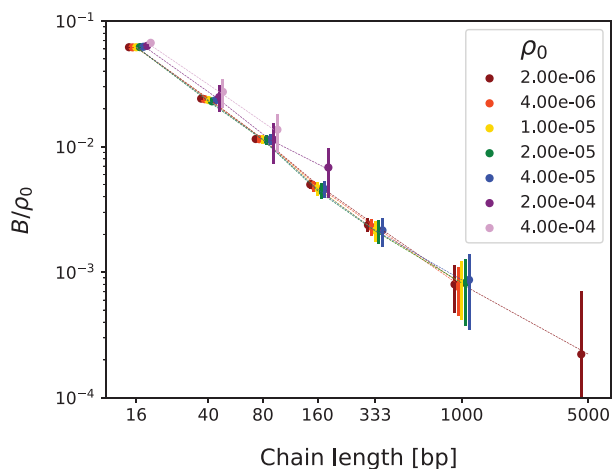
Also of entropic origin is the decreasing dependence of the nematic order modulus  $A$  on the chain length in Figure 4. For a fixed number of segments, stochastic fluctuations of the collective (mesoscopic average) orientational order are smaller for independent segments than for segments bound in a chain. This is simply because the latter have fewer independent orientational degrees of freedom and contribute fewer independent random orientations to the average orientation. As a result, statistical vari-

ations of the average orientation are larger and thus a nonzero nematic order is more probable, which translates into a smaller  $A$ . For longer and hence fewer chains, however, the intra-chain orientational freedom of the segments begins to dominate, which is larger for more flexible chains. A curve  $A(N_s)/\rho_0 \approx 3.60 (1/N_s + 1/82.5)$  fits well the  $A/\rho_0$  data in Figure 4, where the last term is of the order of the inverse persistence length of the chain  $\xi_p$  in units of the segment length, which will be discussed in more detail in Section 4.

The non-monotonic behavior of the elastic constant  $L/\rho_0$ , Figure 4, is rather surprising and remains unaccounted for. It occurs in the regime, where the chain length grows larger than the persistence length, i.e., when energetic nematic elasticity

**Table 1.** Parameters of the model Equation (5) at monomer densities  $\rho_{00}$ ,  $20\rho_{00}$ , and  $100\rho_{00}$ . The values are the same as plotted in Figure 4, the corresponding units are indicated (these are the  $k_B T$ ,  $l_0$  units defined at the end of Section 2). The listed values of the elastic constant, for example, range from  $L = 1.20 \times 10^{-11}$  N (for  $\rho = \rho_0 = 2.5 \times 10^{28} \text{ m}^{-3} \approx 0.050 \text{ mg ml}^{-1}$  and  $N_s = 16$ ) to  $L = 4.32 \times 10^{-9}$  N (for  $\rho = 100\rho_0 \approx 5.0 \text{ mg ml}^{-1}$  and  $N_s = 160$ ).

	$\rho_0 [2.5 \times 10^{28} \text{ m}^{-3}]$	16 bp	40 bp	80 bp	160 bp	333 bp	1000 bp	5000 bp
$A/\rho_0 [4.12 \times 10^{-21}]$	1	0.272	0.129	0.0819	0.0669	0.0574	0.0513	0.0498
	20	0.271	0.129	0.0817	0.0660	0.0569	0.0513	
	100	0.270	0.123	0.0871	0.0727			
$L/\rho_0 [4.76 \times 10^{-40}] \text{ m}^2$	1	1.01	3.30	7.96	6.62	5.68	5.93	6.08
	20	1.04	2.64	6.89	5.84	5.66	5.92	
	100	1.18	3.91	3.62	3.63			
$\tilde{C}/\rho_0 [4.76 \times 10^{-40}] \text{ m}^2$	1	2.73	6.08	13.6	30.3	46.3	64.0	72.2
	20	2.84	7.86	14.8	35.1	52.8	69.6	
	100	3.11	8.62	25.9	40.5			
$B/\rho_0 [4.12 \times 10^{-21}]$	1	0.0620	0.0241	0.0115	0.00499	0.00238	0.000800	0.000222
	20	0.0625	0.0234	0.0114	0.00458	0.00216	0.000869	
	100	0.0635	0.0249	0.0113	0.00683			



**Figure 5.** The log–log plot of  $B/\rho_0$  against chain length  $N_s$  shows a quasi-linear dependence. For better visibility, the points corresponding to different densities are slightly shifted horizontally. Excluding the two highest densities, the slope of the line through the points  $N_s = 16, 40, 80$  is approximately  $-1.17$ .

transitions to entropic elasticity. We should bear in mind, however, that our system is in the isotropic phase, where one cannot expect a well-defined nematic elastic constant.

#### 4. Density–Nematic Coupling Strength

Our prime focus is on the strength of the density–nematic coupling and its dependence on the polymer length. As confirmed by the  $\tilde{G}/\rho_0$  plot in Figure 4, the importance of the coupling, i.e., the relative magnitude of  $\tilde{G}$  with respect to the other model parameters, does not significantly depend on density. In other words, the coupling is already equally effective in the dilute limit and thus ubiquitous for all solutions of DNA, as well as for linear polymeric liquids in general! Moreover, from Table 1 it can be seen that the coupling strength  $\tilde{G}$  is substantially larger than the nematic elastic constant  $L$ , which will prove to be important in the following. The ratio  $\tilde{G}/L$  increases with chain length and already becomes larger than ten for  $N_s = 1000$ . Note that the  $\tilde{G}$  values given represent the actual “measured” strengths of the density–nematic coupling for DNA solution as simulated by the oxDNA model. This is our main empirical result, which fully quantifies the density–nematic coupling. Before comparing it with a theoretical model in Section 4.1, one of our ancillary interests, let us put it in a practical perspective and estimate what it means for the magnitude of the induced nematic order effect. That is, let us connect spatial variations in polymer concentration with the degree of nematic ordering they induce.

We now assume an externally imposed monomer number density profile  $\delta\rho$  depending only on the  $z$  coordinate for simplicity. A nonhomogeneous concentration profile can, for example, be sustained between reservoirs with different chemical potentials. Inclusions such as droplets, bubbles, or impurities – frequently encountered in experimental samples – also act as spatial modulators of the chemical potential. A wave-like, periodic or solitonic concentration profile frequently results spontaneously from a spatial instability in a frustrated system, including phase

separation. It follows from the model Equation (5) with the nonlinearized conservation law Equation (3) for generality (Supporting Information of ref. [16] Section VI, Equation (49))

$$(L + G' \rho^2) \partial_z^2 \delta Q_{zz} + 2G' \rho (\partial_z \delta \rho) \partial_z \delta Q_{zz} + G' (\delta Q_{zz} + \frac{1}{2}) \rho \partial_z^2 \delta \rho - A \delta Q_{zz} = 0 \quad (18)$$

where  $G' \equiv G \left(\frac{2}{3} l_0\right)^2$ . This rather complicated connection between the externally imposed  $\delta\rho(z)$  profile and the resulting uniaxial nematic order profile  $\delta Q_{zz}(z)$  can be reduced to extremely simple expressions in the following special limiting cases.

i) A small amplitude wave-like (sinusoidal) modulation of relative monomer density  $\delta\tilde{\rho} = \delta\rho/\rho_0$  with wave vector  $\mathbf{q} = q\hat{e}_z$  results in the explicit wave-like solution for the induced nematic order,

$$\delta Q_{zz}(z) = -\frac{1}{2} \frac{\tilde{G} q^2}{A + (L + \tilde{G}) q^2} \delta\tilde{\rho}(z) \quad (19)$$

For long-wavelength density modulations, i.e., when the wavelength is large with respect to  $2\pi\sqrt{(L + \tilde{G})/A}$ , which is of the order of the persistence length of the chain, we have  $(L + \tilde{G})q^2 \ll A$  and thus only a small induced nematic order  $\delta Q_{zz} \approx -\frac{1}{2}(\tilde{G}q^2/A) \delta\tilde{\rho}$ . Conversely, for short-wavelength modulations, the induced nematic order is  $\delta Q_{zz} \approx -\frac{1}{2}(1 + L/\tilde{G})^{-1} \delta\tilde{\rho}$  and is of the same order as the relative density variation. The minus sign indicates prolate (axial) ordering of the polymer chain in regions of decreased concentration and oblate (planar) ordering in regions of increased concentration. Note that a perfect prolate order has  $Q_{zz} = 1$  and a perfect oblate order has  $Q_{zz} = -1/2$ .

ii) When a general concentration profile varies slowly on the length scale  $\sqrt{(\tilde{G} + L)/A}$ , we get from Equation (18) in the limit  $\delta Q_{zz} \rightarrow 0$

$$\partial_z^2 Q_{zz}(z) \approx -\frac{1}{2} \frac{1}{1 + L/\tilde{G}} \partial_z^2 \delta\tilde{\rho}(z) \quad (20)$$

That is, a non-uniform relative concentration gradient induces a comparably strong spatial variation of the gradient of the uniaxial nematic order.

We see that the ratio  $L/\tilde{G}$  is the relevant parameter in both cases. If it is small, as we found to be the case in DNA solution, the induced nematic order is substantial – it is of the same order as the relative variation of polymer concentration.

#### 4.1. Comparison With a Theoretical Model for $G$

The observed leading proportionality of  $\tilde{G}$  to the monomer density  $\rho_0$ , i.e., the leading  $1/\rho_0$  dependence of the original parameter  $G$  of Equation (5), is corroborated by the simplest models of the free-energy costs of the sources of Equations (3)–(4) developed in ref. [16]. In these models (see Supporting Information of ref. [16] for derivation), the chain ends are treated as independent vectors (ideal gas of dipolar particles) that contribute to  $\mathbf{g}$ , and the kinks between chain segments as independent contributions to  $\rho_0 l_0 \mathbf{k}$  (recall the definition of both types of sources following

Equation (3)). Chain ends give the entropic free-energy density (Equation (29) of Supporting Information [16]), in the present dimensionless units,

$$\Delta f(g) = \frac{1}{2} \frac{3}{\rho_0^\pm} g^2 = \frac{1}{2} \frac{3N_s}{2\rho_0} g^2 \quad (21)$$

where  $\rho_0^\pm = 2N_c/V = 2\rho_0/N_s$  is the number density of chain ends.<sup>[41]</sup> Kinks between chain segments give the free-energy density (Equation (32) of Supporting Information [16]), in the present dimensionless units,

$$\Delta f(k) = \frac{1}{2} \frac{3\epsilon_T}{2\rho_0} (\rho_0 l_0 k)^2 \quad (22)$$

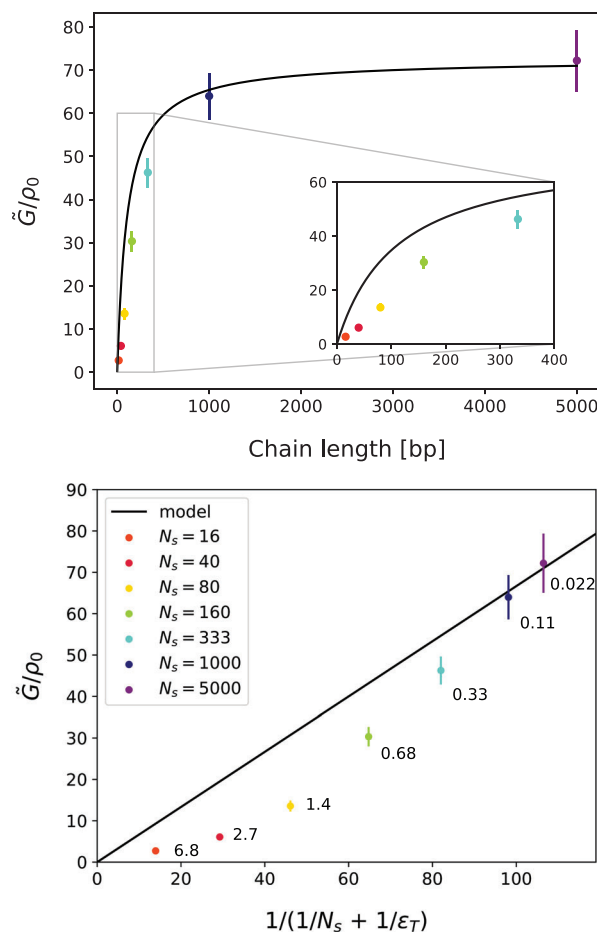
coming from the bending energy of the kinks with dimensionless bending stiffness  $\epsilon_T \equiv \epsilon/(k_b T)$  and bending stiffness  $\epsilon$ . Note that  $\epsilon_T$  is exactly the dimensionless persistence length  $\xi_p$  of the chain,<sup>[42]</sup> p. 399] i.e., the persistence length in units of the segment length  $l_0$ .

We see that both quadratic free energies, Equations (21) and (22) are indeed proportional to  $1/\rho_0$ . However, they are two distinct and practically independent contributions, not a single contribution like the simplistic quadratic penalty potential assumed in the model Equation (5). To get around the inconvenience, we introduce a joint free-energy cost  $\Delta f(h) \equiv \frac{1}{2} Gh^2$  of the combined source  $\mathbf{h} = \mathbf{g} + \rho_0 l_0 \mathbf{k}$ , i.e., a convenient approximation that allows us to live with the single penalty potential of Equation (5). It is obtained by averaging the sum of Equations (21)–(22) over all possible realizations of  $\mathbf{h}$  with respect to  $\mathbf{g}$  and  $\mathbf{k}$  (see Supporting Information of ref. [16] for concept and derivation). The resulting dimensionless strength of the penalty potential is, quite elegantly,<sup>[16]</sup>

$$G = \frac{3}{2} \frac{1}{\rho_0} \left( \frac{1}{N_s} + \frac{1}{\epsilon_T} \right)^{-1} \quad (23)$$

For increasing ratio  $N_s/\epsilon_T$  it shows a crossover from chain-end- to chain-curvature-dominated strength, which is at  $N_s = \epsilon_T$ , i.e., when the persistence length equals the length of the chain. For chains much shorter than the persistence length, we have  $G \approx (3/2)N_s/\rho_0$ , which also follows directly from Equation (21). This is the regime of stiff chains, where the strength of the density–nematic coupling increases proportionally with the number of segments  $N_s$  of the chains. Conversely, for chains much longer than the persistence length,  $G \approx (3/2)\epsilon_T/\rho_0$ , which in turn can be directly deduced from Equation (22). This is the regime of flexible chains, where the coupling strength is controlled by the ratio of bending stiffness and thermal energy. Comparison with  $G \approx (3/2)N_s/\rho_0$  of the stiff regime shows that the persistence length  $\epsilon_T \equiv N_s^{\text{eff}}$  is simply the new effective chain length  $N_s^{\text{eff}}$ , defining the coupling strength in the flexible regime.

To recap, the theoretical coupling strength Equation (23) results from the simplest free-energy models of independent sources of the tensorial conservation law Equation (3), which are additionally approximated as a single source. Does it bear any resemblance to what we measured? The  $\tilde{G}/\rho_0$  plot in Figure 4 indeed shows the cross-over behavior with respect to chain length  $N_s$ . It also shows saturation for large  $N_s$ , which is to be expected according to Equation (23), since in our case the bending stiffness



**Figure 6.** Top: fit of the model Equation (23) with fit parameter  $\epsilon_T$  to the 1000 and 5000 bp points of the  $\tilde{G}/\rho_0$  data from Figure 4;  $\rho_0 = \rho_{00}$ . Bottom: the same data plotted against the expression in Equation (23) (the theoretical curve above is now a straight line). For each data point, the value of the ratio  $\epsilon_T/N_s$  is given.

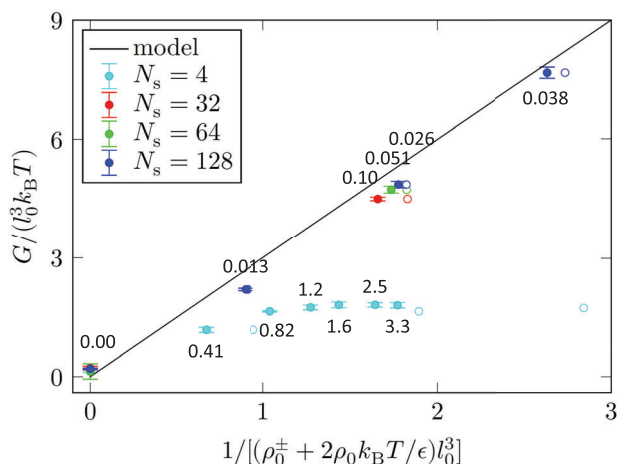
$\epsilon$  of DNA is a fixed parameter and thus its persistence length  $\epsilon_T$  at fixed temperature is also fixed. Thus, we can already confirm at least this qualitative agreement.

Next, we try to fit the measured  $\tilde{G}/\rho_0$  data from Figure 4 with the model Equation (23), i.e., explicitly with

$$\tilde{G}/\rho_0 = \frac{2}{3} \left( \frac{1}{N_s} + \frac{1}{\epsilon_T} \right)^{-1} \quad (24)$$

and the only fit parameter  $\epsilon_T$ . A fit of the  $\rho_0 = \rho_{00}$  data taking into account only points  $N_s = 1000$  and  $N_s = 5000$ , shown in Figure 6 (top), yields  $\epsilon_T \approx 109$ , which is a good estimate of the persistence length (i.e., 109 base pairs). Fits with the last four, three, two, and one of the points  $N_s = \{160, 333, 1000, 5000\}$  all give similar values in the range  $\epsilon_T \approx 102$  to  $\epsilon_T \approx 111$ . Systematic deviations from the theoretical model are observed for shorter chains. In Figure 6 (bottom), the plot of the same data as a function of the expression Equation (24) with  $\epsilon_T = 109$  shows more clearly these discrepancies from the theoretical model Equation (24) (now the straight line). In general, the same value on the horizontal axis can be





**Figure 7.** A plot of the dimensionless coupling strength  $G$  from ref. [16] analogous to Figure 6 (bottom), for a MC-simulated generic isotropic melt of discrete worm-like chains within a mesoscopic “soft” model<sup>[13,33]</sup> (solid circles with error bars). The values of the  $\epsilon_T/N_s$  ratio are again given for all solid circles and can be directly compared with those in Figure 6 (bottom). Empty circles denote abscissae without the  $\rho_0^+$  term, i.e., with neglected theoretical contributions from the chain ends (three rightmost  $N_s = 4$  empty circles lie outside the plot and are not shown). In contrast to Figure 6 only the shortest chains come out of the flexible regime. The permission to reproduce and edit the figure was granted by the author.

realized by different combinations of  $N_s$  and  $\epsilon_T$ . In this particular case,  $\epsilon_T$  is fixed and only  $N_s$  changes – the corresponding ratios  $\epsilon_T/N_s$  are given with the data points.

A worthwhile comparison can be made with an analogous plot in Figure 7, which has been obtained in ref. [16] by MC simulations of a generic isotropic melt of discrete worm-like chains described by a mesoscopic “soft” model.<sup>[13,33]</sup> This is a very different system, but the density–nematic coupling is universal for linear polymers, as is the form of the model Equation (5). In that case, there is data available for four chain lengths  $N_s = 4, 32, 64, 128$  and bending stiffnesses from  $\epsilon_T = 0$  to  $\epsilon_T = 5.0$  for  $N_s = 128$  and up to  $\epsilon_T = 13$  for  $N_s = 4$ , while only  $\epsilon_T = 0$  and 3.3 for  $N_s = 32$  and 64. In Figure 7, the corresponding ratios  $\epsilon_T/N_s$  are again given with the data points. This is a valuable complement to our current data in Figure 6, since the different values on the horizontal axis of Figure 7 are realized predominantly by variations of  $\epsilon_T$  rather than  $N_s$ . Except for the very short  $N_s = 4$  chains, where the theoretical model breaks down due to the small number of intra-chain conformational degrees of freedom, the ratios  $\epsilon_T/N_s$  are small and correspond to the regime of flexible chains. That is,  $\tilde{G}$  is controlled predominantly by  $\epsilon_T$  and is only weakly affected by  $N_s$ , which is also directly confirmed in Figure 7 by the three closely points for equal  $\epsilon_T$  and  $N_s = 32, 64, 128$  with ratios 0.10, 0.051, 0.026, respectively. In Figure 6 (bottom), on the other hand, we gradually cross over to the regime of stiff chains characterized by large  $\epsilon_T/N_s$  ratios. Here, the crossover length (i.e., persistence length,  $\epsilon_T$ ) is much larger than the few segments in the cases of Figure 7, yet the expression Equation (24) already deviates significantly from the measured data (although less severe than the stiffer  $N_s = 4$  chains in Figure 7).

The new data in Figure 6 (bottom) thus indicate that the theoretical model Equations (21)–(24) for the density–nematic cou-

pling strength  $G$  is inaccurate not only for very short chains with few conformational degrees of freedom (segments), but generally also for stiff ( $\epsilon_T/N_s \ll 1$ ) and semi-flexible ( $\epsilon_T/N_s \approx 1$ ) chains, regardless of the number of their segments. Nevertheless, this model is useful also in these regimes and incorporates the influence of the chain ends quite well – without it, the curve in Figure 6 (top) would be just a constant, and the data points in Figure 6 (bottom) would all be on the same(!) abscissa (=109). Taking the chain ends into account clearly corrects the prediction of  $G$  qualitatively in the right direction, and moreover the quantitative inaccuracy for semi-flexible chains is still below 50%.

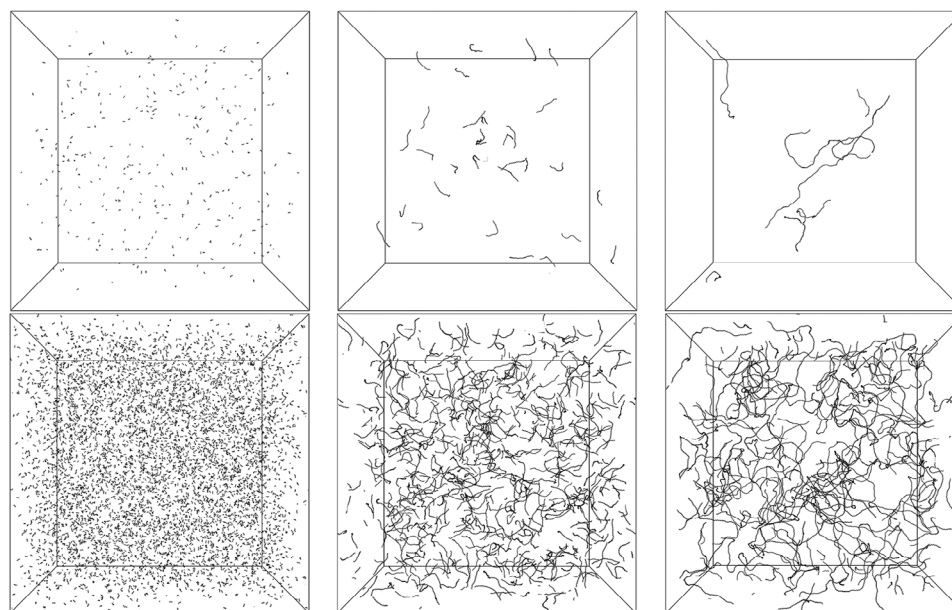
## 5. Conclusion

We have set up a multiscale macroscopic model of an isotropic solution of double-stranded DNA in B-form that comprises the macroscopic polymer concentration field and nematic orientational order field as variables. By conducting microscopic MC simulations employing oxDNA model we determined the parameters of the macroscopic model for various DNA chain lengths from 16 to 5000 base pairs and concentrations from 0.050 mg ml<sup>-1</sup> to 10 mg ml<sup>-1</sup>, which is marginally high for the isotropic phase, at a temperature 25°C and an ionic strength [Na<sup>+</sup>] of 1M. A detailed systematic characterization of this multiscale DNA model, including coverage of temperature and ionic strength ranges, will be carried out in a subsequent study in which macroscopic aspects of the double strand decomposition are of prime interest. Allowing such a decomposition and similar decompositions in other multi-stranded linear polymers, e.g., collagen, requires a generalization of the model for the density–nematic coupling strength.

We found that, to a first approximation, the parameters of the model are simply proportional to the concentration for all concentrations studied, except of the two highest ones. In particular, this means that the coupling between concentration and nematic orientational order, which was our main concern, is already equally effective in the dilute limit and thus inherent to all solutions of DNA (and linear polymer liquids in general), regardless of the concentration. As expected, the coupling strength increases with chain length and saturates for long chains at a value determined by the persistence length.

The results show that for sufficiently long DNA chains, i.e., with a length of at least the persistence length or longer, the density(concentration)–nematic coupling strength  $\tilde{G}$  is significantly larger than the elastic constant  $L$ . Consequently, the induced nematic orientational order is of the same order of magnitude as the relative concentration variation.

To the best of our knowledge, birefringence induced by polymer concentration gradients has not yet been experimentally reported. Nevertheless, due to its fundamental geometrical origin, it should be a very persistent, robust phenomenon, but its large effects may be limited to short length scales supporting large concentration gradients. A large second derivative (curvature) of the concentration field, which is the driving cause of nematic ordering, is typically found in the vicinity of small inclusions of another phase, such as bubbles or droplets. The present work should also serve as motivation for the experimental observation, characterization, and quantification of this interesting phenomenon of



**Figure 8.** Simulation snapshot examples of systems with 16, 160, and 1000 base pairs per chain and base pair densities  $\rho_{00}$  (top) and  $20\rho_{00}$  (bottom).

birefringence due to nematic order induced by spatial variations in polymer concentration.

In addition to the macroscopic aspects of temperature-controlled DNA double strand decomposition and other detailed aspects of isotropic double-stranded DNA solution, our future perspective includes the formulation of its nematic phase analogue, as well as the creation of a variety of multiscale models of other isotropic/nematic linear polymer solutions or melts.

## 6. Methods Section

Simulations were performed at 25°C (298.16K) and an ionic strength  $[\text{Na}^+]$  of 1M. The salt concentration above the typical physiological concentration range 0.01 – 0.1M allowed us to efficiently simulate extensive systems with up to  $1.0 \times 10^5$  base pairs due to a shorter Debye screening length cut-off.

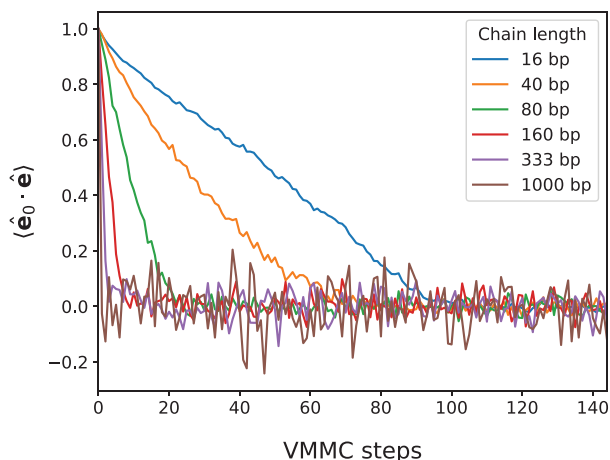
Double-stranded DNA polymers in B-form were constructed using the oxDNA<sup>[22]</sup> generation setup, Figure 2, where each nucleotide base pair is represented as a coarse-grained particle interacting with its environment through backbone connectivity, excluded volume, cross-chain hydrogen bonding and base stacking potentials.<sup>[43]</sup> The B-form, as the by far most common conformation of the DNA macromolecule in aqueous solutions, is excellently represented by the oxDNA coarse-grained implicit solvent model, which enables the exploration of polymer properties in systems comprising 100 000 base pairs or more. For the present purpose, complementary polyguanine and polycytosine chains were simply used, since it had been found that the observed macroscopic properties of the model at the temperature considered did not depend on the specific base pair sequence. Thus, the segments were the guanine-cytosine base pairs with segment length (our length unit)  $l_0 = 0.34 \text{ nm}$ .<sup>[44]</sup>

Generally, the oxDNA model has been shown<sup>[22]</sup> to offer a good representation of structural and mechanical properties, such as persistence length or response to either internal or external

stresses, for various nucleic acid polymers. Moreover, it is particularly well calibrated to reproduce DNA denaturation and hybridization, i.e., dissociation and reassociation of double strand into single strands and vice versa. A separate study is planned to investigate sequence-dependent macroscopic consequences of double strand dissociation, along with their dependence on temperature and ionic strength.

A cubic simulation box with a fixed side of  $L_0 = 1365$  was used (NVT conditions) and periodic boundary conditions were applied. As it was essential for the DNA to be in the isotropic phase, its concentration had to be kept below the level at which cholesteric (chiral nematic) liquid crystalline phase (co-)exists. A systematic array of systems was set up with five different segment (base pair) number densities, multiples 1, 2, 5, 10, and 20 of the base density  $\rho_{00} = 2.0 \times 10^{-6}$  (corresponding to about  $0.050 \text{ mg ml}^{-1}$ ), and monodisperse polymer chains of  $N_s = 16, 40, 80, 160, 333,$  and  $1000$  base pairs at each density, as illustrated in Figure 8. To keep the chain length strictly monodisperse,  $\rho_{00}$  was minimally adjusted as needed by rounding the total number of monomers to the nearest multiple (i.e., integer number of chains  $N_c$ ) of the chain length. This density interval is well below the threshold concentration  $\approx 10 \text{ mg ml}^{-1}$  for the isotropic to cholesteric phase transition of DNA at  $[\text{Na}^+] = 0.1 \text{ M}$ .<sup>[45]</sup> For still higher densities, yet compatible with the isotropic phase, additional systems were set up with  $\rho_0 = 100\rho_{00}$ ,  $N_s = 16, 40, 80, 160$  and  $\rho_0 = 200\rho_{00}$ ,  $N_s = 16, 40, 80$  in a box with  $L_0 = 1365/\sqrt[3]{100}$ , while checking that the degree of nematic order (the average largest in magnitude eigenvalue of  $Q_{ij}$ ) was below  $\approx 0.01$  and the equilibrium phase remained essentially isotropic.

The simulations were performed using the virtual-move Monte Carlo (VMMC) protocol<sup>[43,46]</sup> – a MC algorithm particularly suited for computational models expressed only by pairwise interactions between particles (such as oxDNA). In such an MC step, clusters of interacting particles are moved together while preserving their internal configuration. Under NVT



**Figure 9.** Average projection of the end-to-end directions of the chains (unit vectors  $\hat{e}$ ) onto these directions in the initial configuration (unit vectors  $\hat{e}_0$ ), decreasing with the number of VMMC steps. About 100 steps are sufficient for orientational decorrelation of the shortest chains, while longer chains decorrelate in fewer steps, because they are perturbed more often during a step.

conditions, the VMMC algorithm performs translational and rotational moves of the clusters. The clustering is also subjected to MC evaluation, so strongly bonded particles are only more likely clustered and moved together, while intra-chain MC moves remain present.

The initial configurations, which consisted of parallel DNA chains, were let to equilibrate for at least 2000 VMMC steps, with a large safety margin as confirmed by **Figure 9**. The VMMC trajectories for fluctuation analysis were collected over at least 18 000 VMMC steps. In some cases, more steps were needed to obtain adequate statistics of the density fluctuation – up to 360 000 steps for the shortest chains of 16 base pairs. Standard deviations of the measured fluctuation amplitudes were estimated by block averaging, using a block of size 100.

The fluctuations Equations (13)–(17) of Fourier components of any pair of variables, denoted here as  $\delta a(\mathbf{q}) = \sum_s a_s e^{-i\mathbf{q}\cdot\mathbf{r}^s}$  and  $\delta b(\mathbf{q}) = \sum_s b_s e^{-i\mathbf{q}\cdot\mathbf{r}^s}$ , are extracted as follows

$$\begin{aligned} & \frac{1}{2N_0} [\langle \delta a(\mathbf{q}) \delta b(-\mathbf{q}) \rangle + \langle \delta a(-\mathbf{q}) \delta b(\mathbf{q}) \rangle] \\ &= \frac{1}{N_0} \left\langle \left[ \sum_s a_s \cos(\mathbf{q} \cdot \mathbf{r}^s) \right] \left[ \sum_s b_s \cos(\mathbf{q} \cdot \mathbf{r}^s) \right] \right. \\ & \quad \left. + \left[ \sum_s a_s \sin(\mathbf{q} \cdot \mathbf{r}^s) \right] \left[ \sum_s b_s \sin(\mathbf{q} \cdot \mathbf{r}^s) \right] \right\rangle \end{aligned} \quad (25)$$

where  $s = 1 \dots N_0$  runs over the segments of all chains and  $\mathbf{r}^s$  are their positions. For segment density fluctuations  $\delta\rho$ ,  $a_s = 1$ , and for the nematic fluctuations  $\rho_0 \delta Q_{ij}$ ,  $a_s = (3\mathbf{t}_i^s \mathbf{t}_j^s - \delta_{ij})/2$ , where in B-form DNA the segment directions  $\mathbf{t}_i^s$  are best represented by normals to the nucleobase planes. Note that the low- $q$  components of the extracted discrete variables are by definition “coarse-grained” and hence the long-wavelength correlations Equation (25) computed from the simulation data can be directly

compared to the predictions of the continuum theory Equations (13)–(17).

Since the system is isotropic, all quantities depend only on the magnitude  $|\mathbf{q}| = q$ . This fact was utilized by averaging them over spherical shells with thickness  $\Delta q \sim 2\pi/L_0$ , ensuring that even the smallest shell ( $q \rightarrow 0$ ) with  $q = 0$  excluded was populated. The isotropic symmetry is broken only for non-scalar quantities – by the direction of  $\mathbf{q}$  defining the  $z$  axis, as already in Equations (8)–(12). Thus, expressing their components in such coordinate systems that  $\mathbf{q} = q\hat{e}_z$  for each  $\mathbf{q}$  and, arbitrarily in the  $xy$  plane,

$$\hat{e}_x = \frac{\hat{e}_{x'} - (\hat{e}_{x'} \cdot \hat{e}_z)\hat{e}_z}{|\hat{e}_{x'} - (\hat{e}_{x'} \cdot \hat{e}_z)\hat{e}_z|}, \quad \hat{e}_y = \hat{e}_z \times \hat{e}_x \quad (26)$$

where  $\hat{e}_{x'}$  is aligned with the simulation box, the components are independent of the direction of  $\mathbf{q}$  and can be averaged. With that, for the component  $\rho_0 \delta Q_{zz}$ ,  $a_s$  is  $a_s = [3(\mathbf{t}^s \cdot \hat{e}_z)^2 - 1]/2$  and for the components  $\rho_0 \delta Q_{\{x,y\}z}$ ,  $a_s$  is  $a_s = 3(\mathbf{t}^s \cdot \hat{e}_{\{x,y\}})(\mathbf{t}^s \cdot \hat{e}_z)/2$ .

## Acknowledgements

The authors thank A. Popadić and R. Podgornik for fruitful discussions, help and advice. The authors acknowledge financial support through grants P1-0002, J1-3027, J4-3087 of the Slovenian Research and Innovation Agency.

## Conflict of Interest

The authors declare no conflict of interest.

## Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## Keywords

DNA, density–nematic coupling, isotropic phase, linear polymers

Received: May 24, 2024

Revised: July 26, 2024

Published online:

- [1] P. G. de Gennes, *Mol. Cryst. Liq. Cryst.* **1976**, *34*, 177.
- [2] R. B. Meyer, in *Polym. Liq. Cryst.*, Academic Press, New York **1982**, pp. 133–163.
- [3] G. J. Vroege, T. Odijk, *Macromolecules* **1988**, *21*, 2848.
- [4] R. D. Kamien, D. R. Nelson, *J. Stat. Phys.* **1993**, *71*, 23.
- [5] X. Ao, X. Wen, R. B. Meyer, *Phys. A* **1991**, *176*, 63.
- [6] J. V. Selinger, R. F. Bruinsma, *Phys. Rev. A* **1991**, *43*, 2910.
- [7] D. R. Nelson, *Phys. A* **1991**, *177*, 220.
- [8] P. L. Doussal, D. R. Nelson, *EPL* **1991**, *15*, 161.
- [9] R. D. Kamien, P. Le Doussal, D. R. Nelson, *Phys. Rev. A* **1992**, *45*, 8727.
- [10] J. V. Selinger, R. F. Bruinsma, *J. Phys. II* **1992**, *2*, 1215.
- [11] We are considering generic hairpins of linear chains, not specifically stem-loops<sup>[47]</sup> of nucleic acids, also known as hairpin loops.

- [12] D. Svenšek, G. M. Grason, R. Podgornik, *Phys. Rev. E* **2013**, *88*, 052603.
- [13] A. Popadić, D. Svenšek, R. Podgornik, K. C. Daoulas, M. Praprotnik, *Soft Matter* **2018**, *14*, 5898.
- [14] D. Svenšek, R. Podgornik, *Phys. Rev. E* **2016**, *93*, 052703.
- [15] Note that nematic ordering is fully compatible with hairpins and that, in principle, the degree of nematic order is not affected by hairpins – even in the limit of a large concentration of hairpins, one can have perfect nematic order.
- [16] A. Popadić, D. Svenšek, R. Podgornik, M. Praprotnik, *Adv. Theory Simul.* **2019**, *2*, 1900019.
- [17] The tensorial conservation law is a geometrical analogue of the stationary continuity equation for linear momentum density in an ideal gas system (replacing time with arc length). Following its derivation in [16] (Supporting Information, Section I), Equation (3) can be written as  $\partial_j \tilde{J}_{ij} = \frac{3}{2} g_i + \frac{3}{2} \rho l_0 k_i$ , where  $\tilde{J}_{ij} = \frac{3}{2} \frac{l_0}{V} \sum_n \mu_i^n \mu_j^n$  and the sum runs over all segments in a mesoscopic volume  $V$ . This is to be compared with the stationary linear momentum conservation  $\partial_j \Pi_{ij} = f_i$ , where  $\Pi_{ij} = \frac{m}{V} \sum_n v_i^n v_j^n$  is the mesoscopic momentum flux density of ideal gas particles with masses  $m$  and velocities  $v^n$ .
- [18] M. Praprotnik, L. D. Site, K. Kremer, *Annu. Rev. Phys. Chem.* **2008**, *59*, 545.
- [19] D. Vlachos, A. Mhadeshwar, N. Kaisare, *Comput. Chem. Eng.* **2006**, *30*, 1712.
- [20] J. Fish, G. J. Wagner, S. Keten, *Nat. Mater.* **2021**, *20*, 774.
- [21] Q. Fang, W. Lu, Y. Chen, H. Feng, P. K. Liaw, J. Li, *J. Mech. Phys. Solids* **2022**, *169*, 105067.
- [22] P. Šulc, F. Romano, T. E. Ouldridge, L. Rovigatti, J. P. K. Doye, A. A. Louis, *J. Chem. Phys.* **2012**, *137*, 135101.
- [23] V. G. Taratuta, A. J. Hurd, R. B. Meyer, *Phys. Rev. Lett.* **1985**, *55*, 246.
- [24] X. Ao, X. Wen, R. B. Meyer, *Physica A: Statistical Mechanics and its Applications* **1991**, *176*, 63.
- [25] L. C. A. Groot, M. E. Kuil, J. C. Leyte, J. R. C. V. D. Maarel, R. K. Heenan, S. M. King, G. Jannink, *Liq. Cryst.* **1994**, *17*, 263.
- [26] A. N. Semenov, S. H. Anastasiadis, N. Boudenne, G. Fytas, M. Xenidou, N. Hadjichristidis, *Macromolecules* **1997**, *30*, 6280.
- [27] A. N. Semenov, S. H. Anastasiadis, *Macromolecules* **2000**, *33*, 613.
- [28] K. Binder, S. A. Egorov, A. Milchev, A. Nikoubashman, *J. Phys.: Mater.* **2020**, *3*, 032008.
- [29] M. P. Allen, D. Frenkel, *Phys. Rev. A* **1988**, *37*, 1813.
- [30] M. P. Allen, M. A. Warren, M. R. Wilson, A. Sauron, W. Smith, *J. Chem. Phys.* **1996**, *105*, 2850.
- [31] O. Pelletier, C. Bourgaux, O. Diat, P. Davidson, J. Livage, *Eur. Phys. J. E* **2000**, *2*, 191.
- [32] P. A. O'Brien, M. P. Allen, D. L. Cheung, M. Dennison, A. Masters, *Phys. Rev. E* **2008**, *78*, 051705.
- [33] P. Gemünden, K. C. Daoulas, *Soft Matter* **2015**, *11*, 532.
- [34] S. A. Egorov, A. Milchev, K. Binder, *Phys. Rev. Lett.* **2016**, *116*, 187801.
- [35] A. Milchev, S. A. Egorov, K. Binder, A. Nikoubashman, *J. Chem. Phys.* **2018**, *149*, 174909.
- [36] D. Svenšek, R. Podgornik, *J. Chem. Phys.* **2015**, *143*, 114902.
- [37] Erratum: in Supporting Information of Reference [16] a factor of 1/2 is missing in the first line of Equation (24), i.e., the second quadratic term should be  $\frac{1}{2} |\delta Q_{xx} - \delta Q_{yy}|^2$ . This erratum does not affect the results presented there, since this fluctuation (identical in nature to  $2|\delta Q_{xy}|^2$ ) was not considered.
- [38] K. Nadassy, I. Tomás-Oliveira, I. Alberts, J. Janin, S. J. Wodak, *Nucleic Acids Res.* **2001**, *29*, 3362.
- [39] M. Doi, S. Edwards, *The Theory of Polymer Dynamics*, International series of monographs on physics, Clarendon Press, Oxford **1988**.
- [40] J. Z. Chen, *Prog. Polym. Sci.* **2016**, *54-55*, 3.
- [41] The relation  $\rho_0^\pm = 2\rho_0/N_s$  can be used in our monodisperse case with well-defined number of chain segments  $N_s$ , while  $\rho_0^\pm$  is well-defined also for polydisperse chains.
- [42] L. D. Landau, E. M. Lifshitz, *Statistical Physics*, 3 ed., Butterworth-Heinemann, Oxford, **1980**.
- [43] A. Sengar, T. E. Ouldridge, O. Henrich, L. Rovigatti, P. Šulc, *Front. Mol. Biosci.* **2021**, *8*, 693710.
- [44] W. K. Olson, A. A. Gorin, X.-J. Lu, L. M. Hock, V. B. Zhurkin, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 11163.
- [45] K. Merchant, R. Rill, *Biophys. J.* **1997**, *73*, 3154.
- [46] S. Whitelam, P. L. Geissler, *J. Chem. Phys.* **2007**, *127*, 154101.
- [47] R. A. Cox, *Biochem. J.* **1968**, *106*, 725.