

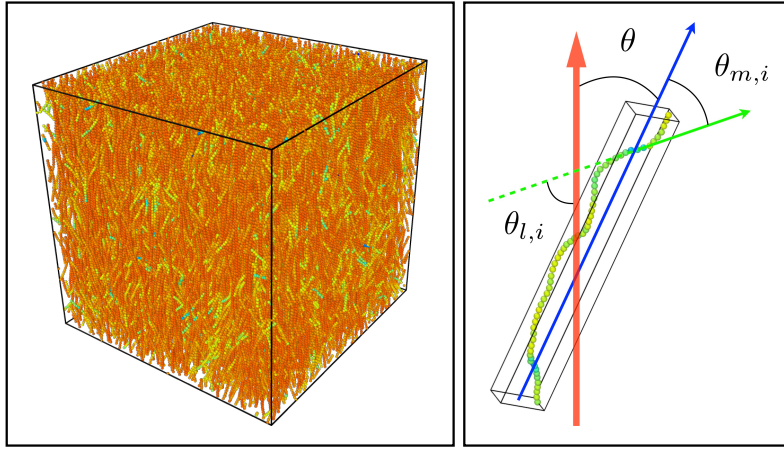
# Incorporating particle flexibility in a density functional description of nematics and cholesterics

Maxime M.C. Tortora and Jonathan P.K. Doye

*Physical and Theoretical Chemistry Laboratory, Department of Chemistry, University of Oxford, South Parks Road, Oxford OX1 3QZ, United Kingdom*

(Dated: 4 May 2018)

We describe a general implementation of the Fynewever-Yethiraj density functional theory (DFT) for the investigation of nematic and cholesteric self-assembly in arbitrary solutions of semi-flexible polymers. The basic assumptions of the theory are discussed in the context of other generalised Onsager descriptions for flexible polyatomic systems. The location of the isotropic-to-nematic phase transition is found to be in good agreement with molecular simulations for elongated chains up to relatively high polymer flexibilities, although the predictions of the theory in the nematic regime lead to gradual underestimations of order parameters with decreasing particle stiffness. This shortcoming is attributed to increasing overestimations of the molecular conformational entropy in higher-density phases, which may not be easily addressed in the formalism of DFT for realistic particle models. Practical consequences of these limitations are illustrated through the application of DFT to systems of near-persistence-length DNA duplexes, whose cholesteric behaviour is found to be strongly contingent on their detailed accessible conformational space in concentrated solutions.



Keywords: Density functional theory; liquid crystals; semi-flexible polymers; DNA.

## I. INTRODUCTION

The self-organisation of polymer solutions into partially-ordered mesophases is a phenomenon of considerable biological relevance, whose occurrences span the formation of phospholipid bilayers for the assembly of cell plasma membranes<sup>1</sup>, the arrangements of F-actin filaments in the cytoplasm<sup>2</sup> and of haemoglobin chains in sickle-cell anaemic blood<sup>3</sup>. This natural ubiquity, combined with a sustained industrial interest in the practical applications of polymeric liquid crystals<sup>4–6</sup>, has spawned a wealth of fundamental and applied investigations of their phase behaviour in a variety of contexts<sup>7,8</sup>.

The first successful mathematical account of emerging order in polymer solutions stems from the seminal work of Onsager<sup>9</sup>, who inferred from first-principle arguments the existence of a concentration-driven transition from a liquid-like isotropic state to an orientationally-ordered nematic state in dispersions of anisotropic colloids. In this framework, which precludes the classical density func-

tional theory (DFT) of non-uniform fluids, the onset of nematic organisation simply results from the competition between orientational entropy and excluded-volume interactions, for which exact explicit expressions may be obtained in the limit of rigid rod-like particles with infinite particle aspect ratios.

Real polymers, however, are generally characterised by a finite flexibility, which may lead to substantial deviations from the slender linear conformations considered by Onsager, and which needs to be accurately incorporated in any reliable statistical-mechanical description of their self-assembly. Considerable efforts have thus been devoted to the generalisation of the Onsager theory to systems of limited stiffness based on various coarse-grained polymer representations and simplified derivations of the corresponding conformational free energy, which provides an additional contribution to the Onsager thermodynamical treatment of their isotropic-to-nematic (I/N) transition<sup>10–24</sup>.

The approximate nature of such approaches therefore generally requires their careful preliminary assess-

ment against the results of molecular simulations in order that the effects of their potential shortcomings may be quantitatively understood, independently from the inevitable ambiguities of direct comparisons with experimental results<sup>25</sup>. However, most current formulations of DFT generally rely on a number of system-specific analytical assumptions and numerical approximations, which do not allow for the straightforward application of theory and simulations to identical particle models. Furthermore, theoretical investigations of the nematic behaviour of semi-flexible chains have so far almost exclusively focused on their uniaxial organisation, while the liquid-crystal phases formed by many biopolymers are commonly observed to rather display a chiral *cholesteric* arrangement, reflecting the detailed symmetries of their local molecular structures<sup>26,27</sup>.

A classical extension of the Onsager theory to the cholesteric assembly of rigid particles was introduced by Straley<sup>28</sup>, who proposed to account for the macroscopic breaking of mirror symmetry in the weak chirality limit based on a microscopic statistical theory of the Frank elastic constants<sup>29</sup>. This perturbative treatment provides a convenient common theoretical formalism in which uniaxial nematic phases are simply assimilated to “degenerate” cholesteric phases with infinite helical pitch, and in which the reliable determination of equilibrium cholesteric structure is contingent on the accurate account of local uniaxial order<sup>28</sup>.

However, in the case of semi-flexible particles, the difficulty of representing the dependence of the conformational free energy on their detailed chiral structure<sup>30</sup> has limited most theoretical investigations of cholesteric ordering to polymer solutions in the coil limit<sup>31,32</sup> — while many experimental cholesteric systems are appreciably stiffer<sup>33</sup>. Furthermore, the complex virial-type coefficients underpinning macroscopic twist in Straley’s approach<sup>28</sup> have only been evaluated analytically for a handful of simplified model systems<sup>30,32,34–38</sup>, and are often used as semi-empirical adjustable parameters when comparing theoretical predictions to experimental cholesteric pitch measurements<sup>33</sup>. Therefore, a general and reliable framework to investigate the link between phase symmetry breaking and molecular chirality in solutions of semi-flexible polymers is still largely lacking.

Building on previous efforts<sup>39,40</sup>, we here propose an efficient implementation of DFT that may be conveniently applied to a wide variety of flexible particle models, taking into account the full details of their microscopic Hamiltonian description. In this hybrid approach, conformational statistics may be introduced through the use of molecular trajectories obtained from direct simulations, thus allowing for a rigorous comparison of theoretical predictions with simulation results. A particular advantage of the method is that it is easily generalisable to Straley’s description of the cholesteric phase, and is suitable for the treatment of arbitrary polymer systems with various degrees of molecular complexity.

The structure of the paper is organised as follows. We

first summarise in Sec. II the main attempts at the inclusion of particle flexibility in the context of DFT, with a particular emphasis on the physical assumptions underlying the different theories, and outline the details of our chosen approach and numerical implementation. We then dedicate Sec. III to the extensive comparison of our results with molecular simulations of the nematic assembly of coarse-grained semi-flexible chains, and present in Sec. IV the application of our method to the cholesteric organisation of near-persistence-length DNA duplexes. Finally, we recapitulate in Sec. V the main conclusions of these analyses, and highlight some potential directions for future research.

## II. DENSITY FUNCTIONAL THEORY FOR SEMI-FLEXIBLE POLYMERS

In the framework of classical DFT, the self-organisation of polyatomic molecules in the absence of external fields is driven by the competition between two coupled contributions, generally referred to as the *ideal* and *excess* free energies  $\mathcal{F}_{\text{id}}$  and  $\mathcal{F}_{\text{exc}}$ , and often defined as the respective intra- and inter-molecular components of their total bulk free energy  $\mathcal{F}$ <sup>41,42</sup>,

$$\mathcal{F} = \mathcal{F}_{\text{id}} + \mathcal{F}_{\text{exc}}. \quad (1)$$

In the so-called Onsager limit of infinitely-stiff particles with high aspect ratios assembling into uniform uniaxial nematic phases, exact analytical expressions for  $\mathcal{F}_{\text{id}}$  and  $\mathcal{F}_{\text{exc}}$  may be derived in the functional form<sup>9</sup>

$$\frac{\beta \mathcal{F}_{\text{id}}[\psi]}{V} = 4\pi^2 \rho \int_0^\pi d\theta \times \sin \theta \psi(\cos \theta) \times \left\{ \log[\rho \psi(\cos \theta)] - 1 \right\}, \quad (2)$$

$$\frac{\beta \mathcal{F}_{\text{exc}}[\psi]}{V} = -\frac{\rho^2}{2} \int_V d\mathbf{r}_{12} \oint d\mathcal{R}_1 d\mathcal{R}_2 \times f(\mathbf{r}_{12}, \mathcal{R}_1, \mathcal{R}_2) \times \psi(\cos \theta_1) \psi(\cos \theta_2), \quad (3)$$

where  $\beta \equiv 1/k_b T$  is the inverse temperature,  $\rho$  the uniform number density and  $\psi(\cos \theta) \equiv \psi(\mathbf{u} \cdot \mathbf{n})$  the orientation distribution function (ODF) describing the degree of alignment of the molecular long axes  $\mathbf{u}$  about the uniform director  $\mathbf{n}$ . In this case, the excess free energy Eq. (3) may fully account for inter-particle interactions at the second-virial level through the Mayer  $f$ -function<sup>43</sup>,

$$f(\mathbf{r}_{12}, \mathcal{R}_1, \mathcal{R}_2) = \exp \left\{ -\beta U(\mathbf{r}_{12}, \mathcal{R}_1, \mathcal{R}_2) \right\} - 1, \quad (4)$$

with  $U(\mathbf{r}_{12}, \mathcal{R}_1, \mathcal{R}_2)$  the interaction energy of two particles with respective orientations  $\mathcal{R}_1, \mathcal{R}_2$  and relative centre-of-mass separation  $\mathbf{r}_{12}$ , while the ideal free energy Eq. (2) simply reduces to that of an ideal monatomic gas of anisotropic particles with ODF  $\psi$ . The equilibrium ODF  $\psi_{\text{eq}}$  is then obtained by functional minimisation of

the corresponding free energy  $\mathcal{F}[\psi]$  at fixed density  $\rho$ ,

$$\left. \frac{\delta \mathcal{F}}{\delta \psi(\cos \theta)} \right|_{\psi_{\text{eq}}} = \rho V k_b T \times 4\pi^2 \lambda \sin \theta, \quad (5)$$

with  $\lambda$  a Lagrange multiplier ensuring the proper normalisation of the ODF. Eqs. (1)–(5) then immediately lead to the well-known self-consistent equation for  $\psi_{\text{eq}}$ ,

$$\psi_{\text{eq}}(\cos \theta) = \exp \left[ \lambda - 1 - \frac{\rho}{4\pi^2} \int_0^\pi d\theta' \times \sin \theta' E(\theta, \theta') \right. \\ \left. \times \psi_{\text{eq}}(\cos \theta') \right], \quad (6)$$

with  $E$  the orientation-dependent generalised excluded volume,

$$E(\theta_1, \theta_2) = - \int_V d\mathbf{r}_{12} \iint_0^{2\pi} d\alpha_1 d\alpha_2 \iint_0^{2\pi} d\phi_1 d\phi_2 \\ \times f(\mathbf{r}_{12}, \mathcal{R}_1, \mathcal{R}_2), \quad (7)$$

where  $\theta_i$ ,  $\phi_i$  and  $\alpha_i$  denote the 3 Euler angles characterising the orientation  $\mathcal{R}_i$  of particle  $i$ . Under these assumptions, the equilibrium structure and thermodynamics of the phase are thus fully determined by Eqs. (1)–(7), which may be readily solved by numerical means for a variety of particle models<sup>39,40</sup>.

However, in the case of particles with finite stiffness, the determination of  $\mathcal{F}_{\text{id}}$  requires the accurate computation of their conformational free energy and its potential dependence on the surrounding medium, which is analytically intractable for all but the simplest model systems<sup>44–46</sup>. Furthermore, the density function  $\psi$  describing the orientational properties of such molecules generally depends on their detailed internal degrees of freedom, rather than their sole long axes  $\mathbf{u}$ , and renders the derivation of their free energy functional  $\mathcal{F}[\psi]$  a complex many-body problem even at the second-virial level. Therefore, several routes have been proposed to construct approximate expressions for  $\mathcal{F}_{\text{id}}$  and  $\mathcal{F}_{\text{exc}}$  in the case of semi-flexible particles, the main features of which we briefly review in the following.

### A. The Khokhlov-Semenov (KS) approach

The earliest analytical attempts can be traced back to the seminal work of Flory<sup>47</sup>, based on a discretised lattice representation of polymer solutions with limited degrees of freedom. However, the many approximations underlying such a basic description have largely restricted its applicability to purely qualitative studies in the context of the I/N phase transition<sup>48</sup>. A more sophisticated treatment was subsequently introduced by Khokhlov and Semenov<sup>10–12</sup>, and further developed by several authors<sup>14–18</sup>, who suggested to combine various extensions of the original Onsager excess free energy

Eq. (3) in the limit of hard rigid cylinders with a mean-field description of several flexible chain models<sup>13</sup>. In this framework, internal particle mechanics thus enter the nematic free energy Eq. (1) purely through the inclusion of a flexibility-dependent orientational entropy term in  $\mathcal{F}_{\text{id}}$ . The resulting expression for  $\mathcal{F}$  in the case of continuous worm-like chains may then be minimised through a self-consistent numerical procedure to determine the contour-dependent equilibrium ODF  $\psi_{\text{eq}}(s, \mathbf{t})$ , quantifying the probability of finding a curvilinear segment  $s \in [0, l_c]$  of the chains with fixed contour length  $l_c$  pointing in direction  $\mathbf{t}$ <sup>18</sup>.

Despite its theoretical elegance and reported successes in the description of the uniaxial nematic properties of fairly-stiff and elongated experimental systems<sup>49</sup>, a significant shortcoming of this approach lies in its strongly coarse-grained nature. Indeed, the KS theory and its extensions effectively reduce the dependence of the phase behaviour of polymers with a given persistence length  $l_p$  to a handful of simplified flexibility mechanisms<sup>50</sup> combined with the two ratios  $l_c/l_p$  and  $l_p/d_{\text{eff}}$ , with  $d_{\text{eff}}$  a somewhat heuristic effective chain diameter subsuming their relevant molecular features<sup>51</sup>. Its generalisation to the treatment of cholesteric order is therefore not straightforward, as the accurate description of phase chirality requires a more detailed account of local microscopic structure<sup>34–38</sup>, and has to our knowledge yet to be convincingly carried out beyond the limiting case of long polymer coils<sup>30–32</sup>, for which  $d_{\text{eff}} \ll l_p \ll l_c$ .

### B. The Fynewever-Yethiraj (FY) approach

A somewhat orthogonal treatment was proposed by Fynewever and Yethiraj<sup>20</sup>, who conversely suggested to account for the effects of flexibility directly through the excess free energy contribution  $\mathcal{F}_{\text{exc}}$ , while retaining the Onsager expression for  $\mathcal{F}_{\text{id}}$  (Eq. (2)). In this context, the virial integral Eq. (3) is instead averaged over a representative ensemble  $\Omega$  of particle conformations as generated by molecular simulations, while the orientational entropy is simply determined by the chain-averaged ODF  $\psi(\mathbf{u} \cdot \mathbf{n})$  describing the molecular ordering of the polymer long axes  $\mathbf{u}$ , as depicted in Fig. 1. The corresponding excess free energy for arbitrary flexible particles may then be written in the form

$$\frac{\beta \mathcal{F}_{\text{exc}}[\psi]}{V} = G(\eta) \frac{\rho^2}{2} \iint_0^\pi d\theta_1 d\theta_2 \times \sin \theta_1 \sin \theta_2 \\ \times \psi(\cos \theta_1) \psi(\cos \theta_2) \times \langle E(\theta_1, \theta_2) \rangle_\Omega, \quad (8)$$

where  $\langle E(\theta_1, \theta_2) \rangle_\Omega$  denotes the ensemble average of the generalised excluded volume Eq. (7) over  $\Omega$ , as discussed in the next paragraphs. Note that following Ref. 20, we have introduced in Eq. (8) the Parsons-Lee (PL) prefactor<sup>52,53</sup>,

$$G(\eta) = \frac{1 - 3\eta/4}{(1 - \eta)^2}, \quad (9)$$

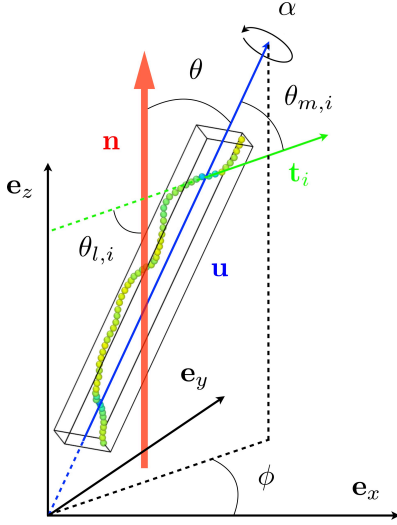


FIG. 1. Particle long axis  $\mathbf{u}$ , local bond orientations  $\mathbf{t}_i$  and nematic director  $\mathbf{n}$  for an arbitrary conformation of the coarse-grained persistent chain model described in Ref. 25. The fixed laboratory frame  $\mathcal{R}_{\text{lab}} \equiv [\mathbf{e}_x \ \mathbf{e}_y \ \mathbf{e}_z]$  is defined such that  $\mathbf{n} \equiv \mathbf{e}_z$ . The black bounding box denotes the eigenvectors of the molecular gyration tensor  $\mathcal{G}$  as determined by the numerical procedure of Sec. II C, and defines the molecular frame  $\mathcal{R}$ .

with  $\eta \equiv \rho v_0$  the volume fraction of the particles with molecular volume  $v_0$ , which provides for a simplified account of the higher-order virial corrections necessary to extend the Onsager description to mesogens with finite aspect ratios. The equilibrium properties of the resulting uniaxial phase may then be determined by solving an extended form of the Onsager self-consistent Eq. (6),

$$\psi_{\text{eq}}(\cos \theta) = \exp \left[ \lambda - 1 - G(\eta) \frac{\rho}{4\pi^2} \int_0^\pi d\theta' \times \sin \theta' \times \langle E(\theta, \theta') \rangle_\Omega \psi_{\text{eq}}(\cos \theta') \right]. \quad (10)$$

In this framework, the inclusion of intra-molecular free energy contributions into Eq. (1) is therefore implicitly relegated to the numerical ensemble-averaged kernel  $\langle E(\theta_1, \theta_2) \rangle_\Omega$ , which circumvents the analytical difficulties involved in the explicit functional derivation of  $\mathcal{F}$  for non-uniform, flexible polyatomic systems<sup>41</sup>. Its results have been found to compare favourably with direct numerical simulations of various athermal solutions of semi-flexible chains, with quantitative agreement being reported in a number of cases<sup>20,54–56</sup>. Furthermore, its relatively simple mathematical formulation is largely independent of the specific features of the molecular system studied, and allows for its straightforward application to a wide range of mesogen models.

This hybrid approach, however, is quite numerically expensive, and suffers from two inherent drawbacks. The first lies in its coarse-grained representation of orientational order at the level of the molecular long axes

$\mathbf{u}$ , as opposed to the more detailed description of local inter-chain correlations between neighbouring monomers through a contour-dependant ODF of the KS-based theories. The second is slightly more subtle, and stems from the tacit assumption in the derivation of Eq. (10) from Eqs. (2), (5) and (8) that the ensemble average  $\langle \cdot \rangle_\Omega$  is independent of the molecular ODF  $\psi$ ,

$$\frac{\delta \langle E \rangle_\Omega}{\delta \psi} = 0,$$

which amounts to neglecting the influence of the variations of the surrounding nematic field on the accessible particle conformational space  $\Omega$ . In the original FY paper,  $\Omega$  was therefore chosen to describe the conformations of a single chain fluctuating in free space — an assumption we refer to as the *unconfined-chain approximation* (UCA) in the rest of the paper.

In order to shed some light on the respective effects of these approximations, and their potential consequences on the capability of DFT to describe realistic polymeric systems, we now outline a generic practical implementation of the FY theory applicable to arbitrary flexible particle models.

### C. Numerical procedure

We have recently introduced a highly-efficient numerical framework based on DFT to work out the emergent properties of uniaxial nematic and cholesteric phases from the structure of their constituent mesogens in the case of arbitrary rigid particle models<sup>39,40</sup>. Its generalisation to the FY treatment of molecular flexibility may be readily performed as follows.

1. Given an arbitrary polyatomic particle model described by a potential  $U_{\text{tot}}$  comprising both bonded and non-bonded interaction forces,

$$U_{\text{tot}} = U_{\text{b}} + U_{\text{nb}},$$

we first construct an ensemble  $\Omega = \Omega[U_{\text{tot}}]$  of uncorrelated particle conformations by means of a single-chain molecular dynamics (MD) or Monte-Carlo (MC) simulation in the canonical ensemble at fixed temperature  $T$ .

2. We then compute the ensemble-averaged excluded volume integral  $\langle E(\theta_1, \theta_2) \rangle_\Omega$  by MC sampling, discarding the bonded components of  $U_{\text{tot}}$ . In the context of Eqs. (4) and (7), a MC step amounts to drawing a random pair of conformations from  $\Omega$ , applying random rigid-body rotations  $\mathcal{R}_1$ ,  $\mathcal{R}_2$  and relative translation  $\mathbf{r}_{12}$  to each respective conformation, and computing the total non-bonded interaction energy  $U_{\text{nb}}$  of the obtained configuration — binning the result for  $E$  using a discrete grid of polar angles  $\theta_i$ .

3. We finally plug the kernel  $\langle E \rangle_\Omega$  into Eq. (10), which may be solved iteratively<sup>57</sup> to determine the equilibrium ODF  $\psi_{\text{eq}}$  at given  $\rho$  and  $T$ .

In the case where the equilibrium phase of the chosen particle model is expected to be cholesteric, rather than uniaxially-symmetric, one may further complement these three steps with the computation of the ensemble-averaged Straley coefficients  $\langle \kappa_{01} \rangle_\Omega$ ,  $\langle \kappa_{11} \rangle_\Omega$ , defined as<sup>28</sup>

$$\beta \kappa_{ij} = -\frac{\rho^2}{2} \int_V d\mathbf{r}_{12} \iint d\mathcal{R}_1 d\mathcal{R}_2 \times f(\mathbf{r}_{12}, \mathcal{R}_1, \mathcal{R}_2) \times \psi_{\text{eq}}^{(i)}(u_1^{(z)}) \left( -u_1^{(y)} r_{12}^{(x)} \right)^i \times \psi_{\text{eq}}^{(j)}(u_2^{(z)}) \left( u_2^{(y)} r_{12}^{(x)} \right)^j, \quad (11)$$

with  $\mathbf{e}_x$  the cholesteric helical axis,  $v^{(k)} \equiv \mathbf{v} \cdot \mathbf{e}_k$  for any vector  $\mathbf{v}$  and  $\psi_{\text{eq}}^{(i)}$  the  $i$ -th derivative of  $\psi_{\text{eq}}$ . Under the assumptions of the FY theory, the quantities  $\langle \kappa_{ij} \rangle_\Omega$  may be conveniently computed following the same method as  $\langle E \rangle_\Omega$ , and the equilibrium pitch  $\mathcal{P}_{\text{eq}}$  of the phase in the limit of weak cholesteric twist is finally given by<sup>40</sup>

$$\mathcal{P}_{\text{eq}}(\rho, T) = 2\pi \times \frac{\langle \kappa_{11} \rangle_\Omega}{\langle \kappa_{01} \rangle_\Omega}. \quad (12)$$

It should be noted that in this perturbative framework, the structure of the cholesteric phase is assumed to be locally indistinguishable from that of the nematic phase, so that all liquid-crystalline properties besides the equilibrium pitch may be identically computed in both cases following the above procedure.

An important prerequisite however lies in the determination of the long axes  $\mathbf{u}$  of the different molecular conformations, with respect to which we choose to quantify the orientational nematic order. Several definitions for such axes have been proposed in the literature in the case of arbitrary-shaped particles, based on the spectral decomposition of the molecular polarisability<sup>58,59</sup>, inertia<sup>58,60</sup> and gyration tensors<sup>61,62</sup>. We here elect to make use of the latter definition, as we feel the purely-geometric character of the gyration tensor  $\mathcal{G}$  to be most consistent with the framework of our equilibrium DFT description, for which molecular electronic and inertial properties are largely irrelevant.

The general expression of  $\mathcal{G}$  for an arbitrary polymer conformation defined by the coordinate matrix  $\mathcal{M} \equiv \{\mathbf{r}_i\}_{i \in [1, N]}$  of its  $N$  constituent monomers reads as

$$\mathcal{G}_{mn} \equiv \frac{1}{N} \sum_{i=1}^N r_i^{(m)} r_i^{(n)} = \frac{1}{N} (\mathcal{M} \cdot \mathcal{M}^T)_{mn} \quad (13)$$

for  $(m, n) \in \{x, y, z\}^2$ , with  $\mathbf{r}_i$  the position vector of the  $i$ -th monomer and  $\mathcal{M}^T$  the matrix transpose of  $\mathcal{M}$ , assuming the origin of the frame to be such that

$$\sum_{i=1}^N \mathbf{r}_i = \mathbf{0}.$$

The particle long axis  $\mathbf{u}$  is then defined as the eigenvector associated with the largest eigenvalue of  $\mathcal{G}$ , and may be efficiently computed from Eq. (13) by principal component analysis (PCA)<sup>40</sup> through the singular-value decomposition of  $\mathcal{M}$ . In the case where all the constituent monomers bear identical masses, it is easy to show that  $\mathbf{u}$  also corresponds to the eigenvector associated with the smallest eigenvalue of the polymer inertia tensor, and one recovers the definition of the molecular axis employed in the original FY implementation<sup>20</sup>.

The use of the UCA at step (1) combined with optimised virial integration methods<sup>40</sup> for the evaluation of Eqs. (7) and (11) enables us to greatly reduce the overall computational expense of the procedure, and thus allows us to tackle flexible mesogen models of arbitrary complexity. We now dedicate the next two sections to the assessment of its results in the case of achiral and chiral systems with various levels of molecular resolution.

### III. THE NEMATIC BEHAVIOUR OF COARSE-GRAINED POLYMER CHAINS

A particularly suitable generic model system for such a quantitative investigation may be provided by the so-called Kremer-Grest (KG) bead-spring chain<sup>63,64</sup>, which constitutes a practical discretised realisation of the original worm-like chain model<sup>65,66</sup> including the effects of both excluded volume and bond flexibility. The bulk properties of their I/N transition were recently investigated in considerable detail by Egorov *et al.*<sup>25</sup>, who combined extensive MD simulations with a simplified implementation of the FY DFT approach to perform a comprehensive mapping of their phase diagram for a wide range of contour lengths, persistence lengths and particle concentrations. The extensive comparison of simulation results with existing theoretical predictions further confirmed that the FY approach consistently outperforms other formulations of DFT in terms of overall agreement with the MD data, despite increasing discrepancies being reported for all theories in the case of highly-flexible chains.

In that study, as in previous applications of the FY theory<sup>20,54</sup>, much emphasis was put on the effects of the PL approximation on the accuracy of the DFT predictions, and several forms of the prefactor Eq. (9) were investigated — though no single expression was conclusively found to yield the best results systematically. Such discussions are undoubtedly highly relevant, especially in the context of the reported shortcomings of the PL rescaling for systems of non-convex particles<sup>39</sup>. However, we here wish to primarily address the much-less documented effects of the coarse-grained description of orientational order and of the UCA, as discussed in Sec. II B, and thus restrict our focus to the regime of fairly-stiff chains with high aspect ratios, for which the PL approximation is expected to work reasonably well, throughout the rest of this section.

In the implementation of the KG model proposed in Ref. 25, polymers are represented as chains of beads interacting through a repulsive Weeks-Chandler-Andersen (WCA) non-bonded potential,

$$u_{ij}^{\text{WCA}}(r_{ij}) = \begin{cases} 4\epsilon \left[ \left( \frac{\sigma}{r_{ij}} \right)^{12} - \left( \frac{\sigma}{r_{ij}} \right)^6 + \frac{1}{4} \right] & \text{if } r_{ij} < r_c, \\ 0 & \text{if } r_{ij} \geq r_c \end{cases},$$

where  $\sigma$  and  $\epsilon$  define the respective model units of length and energy,  $r_c \equiv 2^{1/6}\sigma$  and  $r_{ij}$  denotes the separation distance between any two distinct monomers  $i$  and  $j$ . Particle flexibility is then governed by a simplified bond-bending potential of the form<sup>67</sup>

$$u_i^{\text{bend}}(\theta_i) = \epsilon_b(1 - \cos \theta_i),$$

with  $\epsilon_b$  a bending stiffness parameter and  $\theta_i$  the angle formed by the two bonds linking the consecutive pairs of monomers indexed by  $(i-1, i)$  and  $(i, i+1)$ , respectively. The equilibrium bond length  $l_b \simeq 0.97\sigma$  separating two consecutive monomers finally results from the competition between WCA repulsion and an additional freely-extensible nonlinear elastic bonded potential<sup>63,64</sup>; see Ref. 25 for the details of the full Hamiltonian employed. The persistence length  $l_p$  of the chains is thus fully determined by  $\epsilon_b$ , while their contour length  $l_c$  is simply related to their number  $N$  of constituent monomers through  $l_c \simeq (N-1)l_b$ .

In the framework of Sec. IIB, the Mayer  $f$ -function relative to the pair interaction energy of two arbitrary molecules described by the respective conformations  $(P_1, P_2) \in \Omega^2$  then reads as

$$f(\mathbf{r}_{12}, \mathcal{R}_1, \mathcal{R}_2) = \exp \left( -\beta \sum_{i \in P_1} \sum_{j \in P_2} u_{ij}^{\text{WCA}} \right) - 1,$$

where the double sum runs over all pairs of monomers  $i \in P_1$  and  $j \in P_2$ , and implicitly depends on the centre-of-mass separation vector  $\mathbf{r}_{12}$  and orientations  $\mathcal{R}_1, \mathcal{R}_2$  of the two chains. Following the procedure of Sec. IIC, this quantity may then readily be integrated over the two-particle configurational space and ensemble-averaged over the conformational space  $\Omega$ , where  $\Omega$  has been preliminarily determined by molecular simulations using a given set of parameters for the above potentials. In the following, we choose the model unit of energy such that  $\beta\epsilon = 1$ , and parameterise the Hamiltonian in terms of the chain persistence length  $l_p \simeq \beta\epsilon_b l_b$ <sup>67</sup>, fixing all other constants to the values reported in Ref. 25. We further set the molecular volume  $v_0$  to that of a linear chain of fused hard spheres with radius  $\sigma$  and separation distance  $l_b$ <sup>68</sup>. Note that the latter assumption is somewhat arbitrary; however, the choice of  $v_0$  is only relevant to the PL prefactor (Eq. (9)), whose effects on the I/N transition of the systems considered here are fairly limited, and is of little consequence in our case.

We here choose to evaluate the particle conformational space through basic single-chain MC simulations,

using random translational moves for individual beads combined with a simple Metropolis acceptance criterion. The maximum step size was adjusted over an equilibration run of  $1 \times 10^8$  steps to yield an average acceptance rate of about 30% for a given set of  $l_c$  and  $l_p$ , starting from an initial linear conformation. Production runs of  $2.5 \times 10^9$  steps were performed, and molecular trajectories were constructed by appending one chain snapshot every  $1 \times 10^6$  iterations. The resulting set of 2500 uncorrelated conformations was then used to evaluate the relevant ensemble-averaged integrals in the UCA for each chosen chain stiffness and contour length. This ensemble is henceforth referred to by  $\Omega_0$ , where the subscript denotes the fact that the UCA represents the particle conformational space in the limit of vanishing polymer concentration.

We summarise in Figs. 2 and 3 the comparison of our DFT results with the simulation and theory data of Ref. 25 for chains comprised of  $N = 64$  monomers with  $l_p/l_c \simeq 2$  and 1, respectively. We first remark rather generally that the predictions of our approach in the UCA for the I/N transition density appear to be in good agreement with simulations results in both cases, with the increased particle flexibility of the systems in Fig. 3 deferring the I/N transition to significantly higher volume fractions. However, the further quantitative comparison of nematic properties as described by theory and simulation requires one to carefully distinguish between the different measures of orientational order underlying the FY and KS approaches. In the notation of Sec. II, the dependence of the former's ODF  $\psi_{\text{eq}}^{\text{FY}}$  on the particle long axes  $\mathbf{u}$  allows for the quantification of nematic alignment about the director  $\mathbf{n}$  through the usual uniaxial order parameter  $S$ <sup>26</sup>,

$$S = 2\pi \oint d\mathbf{u} \times \frac{3(\mathbf{u} \cdot \mathbf{n})^2 - 1}{2} \times \psi_{\text{eq}}^{\text{FY}}(\mathbf{u} \cdot \mathbf{n}), \quad (14)$$

while the contour-dependant ODF  $\psi_{\text{eq}}^{\text{KS}}$  of the latter naturally leads to the definition of a generalised uniaxial order parameter  $S_b$ ,

$$S_b = 2\pi \oint d\mathbf{t} \times \frac{3(\mathbf{t} \cdot \mathbf{n})^2 - 1}{2} \times \langle \psi_{\text{eq}}^{\text{KS}} \rangle_c(\mathbf{t} \cdot \mathbf{n}), \quad (15)$$

with  $\langle \psi_{\text{eq}}^{\text{KS}} \rangle_c$  the contour-averaged equilibrium ODF,

$$\langle \psi_{\text{eq}}^{\text{KS}} \rangle_c(\mathbf{t} \cdot \mathbf{n}) = \frac{1}{l_c} \int_0^{l_c} ds \times \psi_{\text{eq}}^{\text{KS}}(s, \mathbf{t} \cdot \mathbf{n}).$$

The two order parameters  $S$  and  $S_b$  therefore generally correspond to rather different physical quantities, as the contour-averaged ODF  $\langle \psi_{\text{eq}}^{\text{KS}} \rangle_c$  appearing in Eq. (15) effectively incorporates the intra-molecular dispersion of the chain segment orientations  $\mathbf{t}$ , which is not directly accounted for by the coarser-grained  $\psi_{\text{eq}}^{\text{FY}}$  of Eq. (14). It is proven in Appendix A that under the assumptions of the FY theory, the general relation between  $S$  and  $S_b$  may be derived in the compact form

$$S_b = S \times S_{\text{int}}, \quad (16)$$



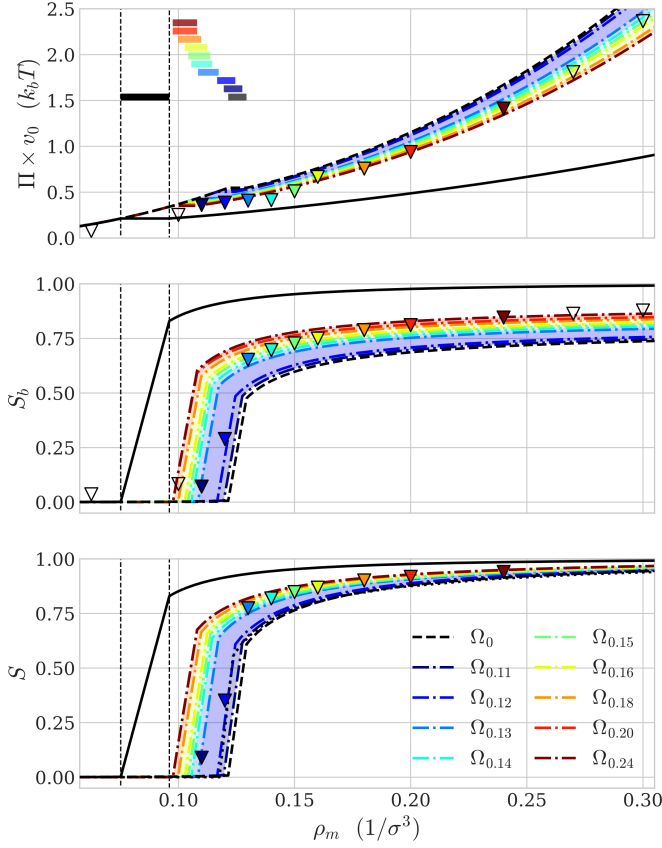


FIG. 2. Osmotic pressure  $\Pi$ , bond order parameter  $S_b$  and molecular order parameter  $S$  as a function of monomer number density  $\rho_m$  for KG chains with persistence length  $l_p = 128 l_b$  and  $N = 64$  beads (contour length  $l_c = 63 l_b$ ). Markers denote the simulation results of Ref. 25, from which  $S$  and  $S_b$  were determined through standard means as the largest eigenvalues of the corresponding  $Q$ -tensors<sup>58</sup>. Black solid lines represent our DFT predictions for fully-rigid chains ( $\Omega_\infty$ ). Black dashed lines correspond to our DFT results in the UCA ( $\Omega_0$ ), and the black dotted line denotes the DFT data of Ref. 25 (only for  $S$ ). Dash-dotted lines were computed using the conformational space  $\Omega_\rho$  obtained from the bulk simulations of Ref. 25 at various fixed densities  $\rho$ , as described in the text. The coloured stripes in the upper panel delimit the respective I/N coexistence ranges predicted by the FY theory using the different conformational spaces, in which order parameters were computed through a standard lever-rule interpolation between the isotropic and nematic binodal points.

with  $S_{\text{int}}$  an ensemble-averaged intra-molecular order parameter describing the distribution of chain segments  $\mathbf{t}$  about the long axes  $\mathbf{u}$  of the polymers as characterised by an arbitrary conformational space  $\Omega$ ,

$$S_{\text{int}} = \left\langle \left\langle \frac{3(\mathbf{t} \cdot \mathbf{u})^2 - 1}{2} \right\rangle_c \right\rangle_\Omega. \quad (17)$$

In the following, we assimilate the generic chain segment  $\mathbf{t}(s)$  to the normalised bond vector  $\mathbf{t}_i$  joining the monomers  $i$  and  $i + 1$  of a discrete KG chain, and thus

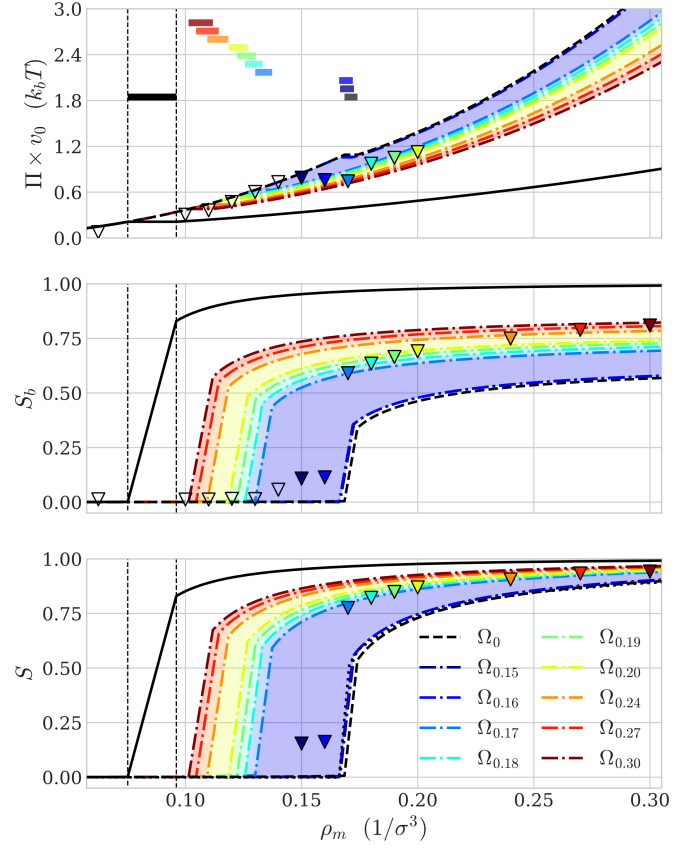


FIG. 3. Same as in Fig. 2 for chains with  $l_p = 64 l_b$ .

refer to  $S$  and  $S_b$  as the *molecular* and *bond order parameters*, respectively.

The comparison of simulation results with DFT in the UCA in terms of  $S$  reveals that DFT tends to gradually underestimate the level of alignment in the nematic phase with increasing particle flexibilities, contrary to the findings of Ref. 25. This discrepancy stems from the fact that the authors of that study actually compared the bond order parameter  $S_b$  calculated from simulations with the molecular order parameter  $S$  of the FY description. Our predicted values for  $S$  are however found to be virtually identical to the DFT data reported in Ref. 25, in which a simplified implementation of the FY theory was used based on an empirical expression for  $\langle E \rangle_{\Omega_0}$  reported in Ref. 20 for semi-flexible tangent hard-sphere chains. This observation demonstrates that the latter system provides a reliable approximation of the KG chains in this context.

The DFT underestimation of nematic order can be easily understood as a direct consequence of the UCA, which does not account for the potential stretching of the chains due to the surrounding nematic field, as the particles gradually forego conformational entropy to optimise their liquid-crystalline packing<sup>14</sup>. It was thus reported in Ref. 25 that while the root-mean-square end-to-end distance  $L(\rho)$  generally remains roughly constant in the isotropic phase, the I/N transition is usually associated with a rather steep jump in  $L$  followed by a

slower increase towards the saturation value  $l_c$ , as the chains progressively straighten along the nematic director. Therefore, the UCA remains quantitatively accurate up to the I/N transition, but increasingly overestimates the effects of particle flexibility as one moves deeper into the nematic range — especially for chains with lower relative stiffnesses  $l_p/l_c$ , whose accessible conformations are most affected by these phase-induced restrictions<sup>25</sup>.

The effects of this shortcoming of the FY theory may be probed by replacing the previous single-chain conformational space  $\Omega_0$  obtained in the limit of infinite dilution with an ensemble  $\Omega_\rho$  of single-chain conformations obtained from bulk molecular simulations at fixed finite density  $\rho$ . We here use for  $\Omega_\rho \equiv \Omega_{\rho_m}$  a set of 30 000 particle conformations obtained from the analysis of the simulation data of Ref. 25 at various monomer densities  $\rho_m$ , as indicated in Figs. 2 and 3. As expected from the previous discussion, we thus find the DFT results for  $\Omega_\rho$  taken at simulation state points in the isotropic phase to be in very good agreement with those obtained using  $\Omega_0$  in the UCA, which indicates the absence of significant chain stretching prior to the I/N transition for the systems considered here.

While this observation confirms the suitability of the UCA to describe the free energy of the isotropic phase in our case, the use of  $\Omega_0$  in the nematic phase leads to an overestimation of the particle conformational entropy associated with an underestimation of their collective packing entropy, resulting in the observed underestimation of local nematic order. The partial compensation of these two effects in the total nematic free energy  $\mathcal{F}$  may thus explain the ability of the UCA to correctly capture the rough location of the I/N transition zone, although its detailed predictions in terms of I/N coexistence densities should generally be interpreted with caution, as discussed in the next paragraph. Conversely, the use of simulated ensembles  $\Omega_\rho$  obtained at nematic state points leads to increasing overestimations of chain stretching in the lower-density regions, and thus gradually pushes the I/N transition predicted by DFT to unrealistically-low concentrations for the more flexible particles.

Hence, it is apparent from Figs. 2 and 3 that while the use of a given ensemble  $\Omega_\rho$  generally enables DFT to correctly reproduce simulation results at that same density  $\rho$ , owing to the state-function character of the underlying free energy  $\mathcal{F}$ , no single representation of the particle conformational space may yield an accurate description of thermodynamic properties at all concentrations in the framework of the FY description. Such limitations are particularly stringent for the accurate computation of the I/N coexistence range due to the abrupt density variations of the conformational space in this region, illustrated for instance by the wide gap between the DFT predictions relative to the pre-transition  $\Omega_{0,16}$  and the post-transition  $\Omega_{0,17}$  in Fig. 3. These effects need to be taken into account for the reliable resolution of the coupled phase-coexistence equations, and may explain the extremely-narrow width of the coexistence domains pre-

dicted by DFT in Figs. 2 and 3, delimited by the plateaus in the osmotic pressure  $\Pi$ . However, the lack of detailed I/N coexistence analysis in the simulations of Ref. 25 precludes the thorough quantitative investigation of these considerations, whose effects on the comparison of DFT results with experimental observations are further discussed in Sec. IV.

Finally, one may investigate the impact of the coarse-grained treatment of orientational order through the calculation of the DFT-predicted bond order parameters  $S_b$  from Eqs. (16) and (17). Comparison with simulations suggests that the underestimations of  $S_b$  using the UCA are more pronounced than for  $S$ , which could be partly attributed to the unsuitability of the FY molecular ODF to describe the local ordering of neighbouring chain segments in solution. However, the good quantitative agreement for  $S_b$  achieved at given density  $\rho$  through the use of  $\Omega_\rho$  suggests that the main limiting factor may actually lie in the UCA itself. Indeed, it can be seen from Eq. (16) that

$$S_b \leq S_{\text{int}},$$

with the equality being attained in the limit of perfect crystalline order for the particle long axes. It is shown in Appendix B that for fairly-stiff unconfined chains, the intra-molecular order parameter  $S_{\text{int}}$  may be approximated as

$$S_{\text{int}} \simeq 1 - \frac{l_c}{2l_p}.$$

The slow saturation of  $S_{\text{int}}$  with increasing persistence lengths, which reflects the presence of small undulations of the free chains about their long axes, therefore imposes an unrealistic upper bound for  $S_b$  in the UCA even in the case where  $l_p$  is significantly larger than  $l_c$ . One thus recovers the important concept of *deflection length*<sup>14</sup>, denoted by  $l_d$ , which quantifies the average length-scale over which polymers are diverted from their unconfined conformations by the presence of the surrounding nematic field. While our results suggest that the FY molecular order parameter  $S$  may be reasonably captured by the UCA for chains such that  $l_p \gtrsim 2l_c$ , the accurate description of the KS bond order parameter  $S_b$  in the UCA thus requires the much more stringent inequalities

$$l_c < l_d \ll l_p,$$

which corresponds to the so-called *rigid-rod limit* where the effects of particle flexibility may be neglected altogether<sup>19</sup>. This limitation therefore requires the description of the detailed conformational statistics of polymers in a confining nematic field as a necessary prerequisite for any form of DFT based on a finer-grained representation of orientational order. This task may be further complicated by the reported importance of collective deflection fluctuation modes<sup>25,69</sup>, which may not be easily captured by the mean-field treatment of KS-based descriptions, and may entail significant further theoretical developments for the reliable implementation of such approaches.



#### IV. THE CHOLESTERIC BEHAVIOUR OF DNA DUPLEXES

As an example application of the techniques introduced here to a more complex and experimentally-realistic system, we now turn our focus to solutions of DNA duplexes. DNA constitutes a well-studied semi-flexible biopolymer with relatively-high stiffness, whose persistence length has been measured to be 130 base pairs (44 nm) at salt concentration  $c_{\text{Na}^+} = 0.5 \text{ M}$ <sup>70</sup>, although the precise value is somewhat dependent on solution conditions<sup>70,71</sup> and sequence<sup>72</sup>. DNA is also chiral, both in terms of its excluded volume — it is a grooved double helix with the major groove larger than the minor groove — and its electrostatics — it has a double helical pattern of negative charges associated with the phosphate groups along the backbone. Consequently, DNA duplexes in concentrated solutions generally assemble into cholesteric phases, which exhibit a rich structural polymorphism<sup>73–76</sup>. Highly-oriented dense DNA packings are also relevant to biology, for example, occurring in bacterial nucleoids<sup>77</sup>, dinoflagellate chromosomes<sup>78</sup> and sperm cells<sup>79</sup>.

Here we will focus on DNA duplexes that are of the order of the persistence length. In particular, we will compare to the results of Refs. 80 and 81, with respectively consider DNA duplexes with reported most-probable lengths of about 146 base pairs. As the diameter of DNA duplexes due to excluded volume is about 25 Å, these examples have an aspect ratio of about 20 — although their effective aspect ratio in solution is expected to be somewhat lower due to the effects of electrostatics<sup>9,82</sup>. The former study focuses on the phase diagram, in particular the coexisting concentrations of the cholesteric and isotropic phases<sup>80</sup>, whereas the latter’s emphasis is on the dependence of the cholesteric pitch on DNA concentration and solution conditions<sup>81</sup>. One of the particularly noteworthy results of the latter is that the cholesteric phase is left-handed with a pitch that depends non-monotonically on the DNA concentration.

The left-handed character is intriguing, as basic considerations based on the packing of hard threaded objects suggest that right-handed particles with a thread angle less than 45° (it is about 30° for DNA) should exhibit a right-handed phase<sup>28</sup>. This simple argument has been found to be reasonably well-obeyed for model systems with purely steric interactions<sup>40,83</sup>. However, the addition of electrostatics may potentially add an additional level of complexity. If one considers the negative charges of the phosphates only, it has been postulated that the electrostatic repulsion would be minimised in a left-handed configuration<sup>84,85</sup>, although recent simulation results for coarse-grained particles with helical charge distributions appear to challenge this hypothesis<sup>86–88</sup>. Conversely, this preference has also been suggested to be reversed if one further considers the possibility of counterion condensation within the grooves of the helix<sup>85</sup>.

To try to get a more rigorous insight into these ques-

tions, a number of theoretical studies have attempted to explain the observed experimental behaviour. Firstly, Kornyshev *et al.* exclusively considered the potential chirality arising from electrostatic interactions, and although they predicted a dependence of the pitch on concentration that resembled experiment<sup>89</sup>, it was later revealed that the corresponding pitches had in fact the wrong handedness<sup>85</sup>. Secondly, Tombolato and Ferrarini applied a DFT approach to a simplified representation of the DNA duplexes, and found that when only the steric interactions were considered, a right-handed phase resulted — in agreement with the previous geometric argument. By contrast, when electrostatic interactions were included, this tendency was reversed and a left-handed phase was predicted, whose pitch was further reported to be in reasonable agreement with experiment for the single DNA concentration considered<sup>84</sup>.

However, numerous approximations were introduced in this study for the computation of the virial-type integrals in Eq. (11), which underpin the determination of the equilibrium pitch  $\mathcal{P}_{\text{eq}}$  from Eq. (12) in the framework of Straley’s DFT<sup>28</sup>. As phase chirality is generally very weak in experimental cholesterics, DFT predictions for the pitch of such systems can be very sensitive to the accuracy of the calculations of such integrals. Thus, we have previously shown<sup>39</sup> that such results should generally be interpreted with caution, unless a well-defined numerical approach is used that allows for the demonstrable convergence of the procedure to the required accuracy.

Both of the above studies also treated DNA as an infinitely rigid molecule. However, flexibility is likely to have a large impact on the phase behaviour of polymers whose contour lengths are of the order of the persistence length, as is evident from Sec. III. The potential effects of flexibility on the cholesteric pitch are largely unexplored<sup>88</sup>, but should be expected to be equally significant in light of the previous discussions.

Here, we apply the current DFT approach to 146-base pair DNA duplexes. In order to both sample the particle conformational space and to calculate the inter-duplex interactions in the procedure of Sec. II C, we make use of the oxDNA model<sup>90,91</sup>. OxDNA is a nucleotide-level coarse-grained model that has been widely applied to the study of DNA-based systems, both in the context of biophysics and nanotechnology<sup>92</sup>. It provides a correct description of the mechanical properties of double-stranded DNA<sup>93,94</sup>, exhibits sensible values for the persistence length and the twist modulus<sup>91</sup>, and has been shown to also capture more subtle phenomena such as twist-bend coupling<sup>95</sup>. We note that the electrostatic interactions in the model are of a simple Debye-Hückel-like form and have been parameterised to reproduce the salt-concentration dependence of the thermodynamics of hybridisation<sup>91</sup>, rather than any specific properties related to inter-duplex interactions.

We now generate the UCA ensemble  $\Omega_0$  by means of single-duplex MD simulations, using an Anderson-like thermostat at temperature  $T = 20^\circ \text{C}$ . Equilibration was

TABLE I. Isotropic/cholesteric coexistence concentrations for near-persistence-length B-DNA duplexes at  $T = 20^\circ$  and various salt concentrations.  $\Omega_0$  and  $\Omega_\infty$  denote the DFT predictions of this work obtained in the UCA and in the limit of infinite duplex stiffness, respectively, and are compared with the experimental results of Refs. 80 and 81. All values are reported in mg/ml.

$c_{\text{Na}^+}$	Ref. 80		Ref. 81		$\Omega_\infty$		$\Omega_0$	
	$c_i$	$c_c$	$c_i$	$c_c$	$c_i$	$c_c$	$c_i$	$c_c$
0.5 M	n/a	n/a	n/a	218	120	138	242	250
1.0 M	171	270	n/a	257	132	155	268	276

performed over  $1 \times 10^7$  simulation steps, and production runs of  $1 \times 10^9$  steps were used to generate a set of 2000 uncorrelated configurations. The molecular volume  $v_0$  relative to the PL approximation (Eq. (9)) was computed using the average B-DNA nucleotide volumes measured in Ref. 96, and mass concentrations were obtained assuming a molar weight of 650 Da per base pair.

The results for the isotropic-to-cholesteric (I/C) transition densities are presented in Table I. As expected from Sec. III, the onset of cholesteric order occurs at a much lower concentration when DNA is considered to be rigid. Significantly, the predictions when flexibility is taken into account in the UCA are in much better agreement with the experimental densities  $c_c$  reported in Refs. 80 and 81, although the latter's measurements may not very accurately reflect the true cholesteric binodal point<sup>81</sup>. The dependence of the results on salt concentration is also found to follow the experimental trends, with stronger electrostatic screening generally deferring the I/C transition range to higher densities, as predicted by theory<sup>82</sup>.

One notable difference with the results of Ref. 80 lies in the width of the coexistence region, which is found to be strongly underestimated by DFT. This discrepancy may be partially attributable to the shortcomings of the UCA in the vicinity of the I/C transition, as discussed in Sec. III, and could also reflect potential inaccuracies of the oxDNA model. Furthermore, our theory does not account for the substantial polydispersity reported in the experimental systems<sup>80</sup>, which should be expected to significantly widen the phase coexistence range<sup>49,51</sup>.

We now turn our focus to the cholesteric pitch, and first consider rigid straight duplexes in order to compare our results with those previous theoretical investigations. As illustrated in Fig. 4, we find the predicted pitch to be right-handed when we do not include the electrostatic term in the potential, as expected from the previous excluded volume argument. Interestingly, the magnitude of the pitch is also found to be much larger than the experimental values, indicating a significant underestimation of phase chirality. However, unlike in Ref. 84, we do not report evidence of a handedness inversion when electrostatic repulsion is taken into account, but rather observe a further reduction of the predicted phase chirality.

This observation could potentially be explained by the

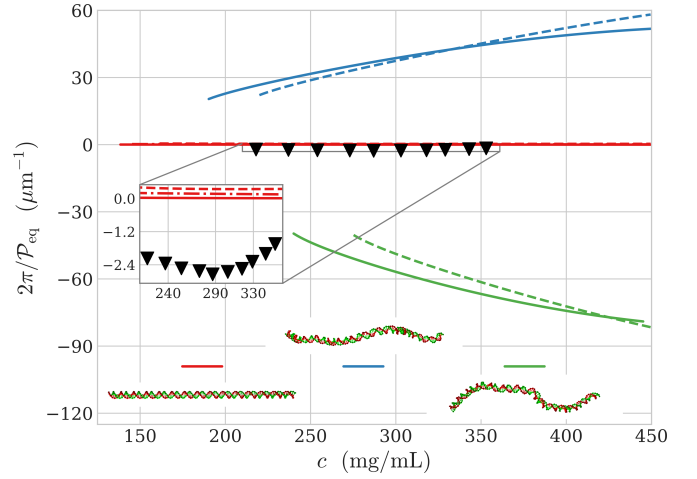


FIG. 4. Inverse equilibrium cholesteric pitch  $2\pi/P_{\text{eq}}$  as predicted by DFT for 146 base-pair DNA duplexes in the limit of infinite particle stiffness ( $\Omega_\infty$ , in red) and for two instantaneous conformations, respectively observed to possess a strong left- (blue) and right-handed (green) character. Solid lines were computed using the oxDNA-parameterised excluded volume and Debye-Hückel interactions at  $T = 20^\circ$  and  $c_{\text{Na}^+} = 0.5$  M. The dash-dotted line was similarly calculated by replacing the Debye-Hückel contribution with the electrostatic potential of Ref. 84 (only for  $\Omega_\infty$ ). Dashed lines were obtained in the limit of pure steric repulsion ( $c_{\text{Na}^+} \rightarrow \infty$ ). Markers denote the experimental results of Ref. 81. The lowest densities reported correspond to the cholesteric binodal points  $c_c$  predicted by DFT.

competition between the entropic favourability of right-handed duplex configurations due to excluded-volume contributions and the hypothetical enthalpic preference for left-handed arrangements to minimise electrostatic repulsion, as postulated in Ref. 84. However, we find a more likely explanation to lie in the fact that electrostatics penalise the close-pair configurations where the chirality of the excluded volume is most relevant, so that longer-range repulsion simply ‘smears out’ the local details of the chiral surface structure of DNA. To verify that this effect is not simply an artefact of our chosen description of electrostatic contributions in the oxDNA model, we repeated the calculations using the electrostatic potential detailed in Ref. 84. Our results were however found to be very similar in both cases, which leads to the conclusion that the prediction of a left-handed phase in Ref. 84 is most likely due to some of the numerical approximations introduced in the virial calculations therein.

We have also attempted to compute the cholesteric pitch including the effects of flexibility in the UCA. However, we concluded after extensive calculations that the drastic dependence of the predicted pitches on the sampling of the particle conformational space precluded their determination with sufficient statistical accuracy. To illustrate the nature of the problems, we report in Fig. 4 the equilibrium pitches computed for two example con-

formations taken from  $\Omega_0$ , both of which were observed to possess noticeably helicoidal shapes with opposite respective handednesses. Interestingly, the handedness of the equilibrium pitch is then obtained to be opposite to that of the corresponding helicoidal conformation, which mirrors theoretical predictions for hard, weakly-curved coarse-grained helices<sup>39,83,97</sup>. Furthermore, the phase chirality is found to be considerably stronger for these conformations, and the resulting pitches only weakly depend on the inclusion of electrostatics, as the shape helicity occurs on much longer length-scales than that of the screened electrostatic interactions.

The strong variations in the predicted pitches between individual DNA conformations prevented the reliable convergence of the DFT procedure, and the large fluctuations in both their handedness and magnitude obtained using different independent ensembles  $\Omega_0$  did not allow for the observation of any meaningful tendencies. Therefore, we could not reliably determine whether such helical conformations may be linked to the underlying intramolecular chirality of DNA, nor what their net effect on the cholesteric pitch should be, given a large enough statistical sample of the conformational space. Moreover, it is apparent from Sec. III that the UCA should generally be unsuitable for the description of the nematic — and *a fortiori* cholesteric — phases of particles whose contour lengths are of the order of the persistence length, so we made no attempt to pursue this study further.

It nonetheless transpires that the flexibility of DNA duplexes plays a significant role in their liquid-crystalline assembly, which cannot be reasonably neglected in any theoretical investigation of their phase behaviour — especially in the context of the fine determination of their cholesteric pitch. However, the quantitative examination of its effects is an arduous task, as near-persistence-length duplexes are too flexible to be accurately described by the FY DFT in the UCA, while the aspect ratios of their shorter counterparts may be too low for the justifiable use of the PL approximation and other empirical extensions of the Onsager DFT.

In the current framework, one possible way to circumvent these issues would of course be to use the oxDNA model to simulate a nematic phase of elongated DNA duplexes at the concentration of interest, in order to generate ensembles of configurations that are more representative of a real cholesteric phase to be input into the FY DFT. However, such simulations may prove to be quite computationally challenging, although they could provide an interesting avenue for future research.

## V. CONCLUSION

We have investigated in detail the application of the FY DFT to the description of the nematic self-assembly of various semi-flexible particle models, and showed for the first time how it can be straightforwardly extended to account for the effects of particle flexibility on the

cholesteric phase. Extensive comparisons with the results of molecular simulations of persistent KG chains reveal that the I/N transition densities of polymers with  $l_p/l_c \gtrsim 1$  appear to be well captured by the theory in the UCA, although the detailed mechanisms underlying these predictions are not fully clear. This limitation stems from the fact that while the FY DFT is found to be quantitatively accurate in the isotropic phase, where no concentration-dependent stretching of the chains is observed for the systems considered here, the reliable computation of nematic properties beyond the I/N transition generally requires a more involved account of the restrictions of the accessible particle conformation space induced by the confining effects of the surrounding nematic field.

These considerations deserve to be investigated further, and dedicated numerical studies of the I/N coexistence properties of persistent chains, along the lines of the simulations of Ref. 98 for systems of jointed spherocylinders, would be desirable. Moreover, the effects of these restrictions should generally depend on the scaled contour length  $l_c/\sigma$  as well as  $l_p/l_c$ , while the former ratio was kept fixed in this paper. However, the quantitative agreement between FY DFT predictions and simulation results reported in Ref. 25 was found to be relatively unaltered by varying of the contour length in the range  $l_c/\sigma \in [15, 63]$  at given  $l_p/l_c$ , so it is expected that the previous discussions should quite generally apply to elongated semi-flexible particles.

We additionally demonstrate that the use of a molecular-level representation of orientational order allows the FY theory to circumvent some of these shortcomings by coarse-graining over conformational fluctuations on the scale of the deflection length, which may provide for a satisfactory description of the nematic phase of polymer chains with  $l_p/l_c \gtrsim 2$ . While this statement may be somewhat limiting in scope, we feel that such a restriction is largely inherent to the underlying formalism of the Onsager DFT, as the complex coupling between conformational statistics and nematic order is further supplemented by the growing inadequacy of the second-virial approximation to describe the higher-density I/N transitions associated with increased particle flexibilities. It is also worth remarking that previous numerical studies have demonstrated the quantitative accuracy of the FY approach to consistently match (and in many cases surpass) that of other DFT implementations across a wide range of chain contour lengths  $l_c$  and relative persistence lengths  $l_p/l_c$ <sup>20,21,25</sup>, which points to the rather general unsuitability of existing theoretical frameworks to tackle polymeric systems of intermediate stiffness<sup>49</sup>.

Such limitations are particularly stringent for the determination of cholesteric properties, which are especially sensitive to the detailed nature of the conformational fluctuations undergone by flexible particles. It is thus found that while the combination of the FY DFT and the oxDNA coarse-grained model provides for a satisfactory prediction of the I/N transition of near-persistence-

length DNA duplexes, the strong dependence of the predicted pitches on minute structural changes largely precludes their realistic determination in the context of the UCA.

This analysis nonetheless reveals that steric and electrostatic interactions between rigid linear DNA duplexes appear to be unable to account for either the handedness or the magnitude of the phase chirality observed in experiments. This conclusion concurs with the recent findings of Ref. 99, in which extensive all-atom MD simulations of short DNA oligomers demonstrated the absence of a significant chiral contribution attributable to inter-duplex electrostatic interactions — thus suggesting the need for alternative mechanisms to explain their cholesteric behaviour.

The accurate description of the conformational statistics of DNA duplexes in dense phases therefore emerges as a crucial prerequisite to any realistic investigation of their cholesteric assembly, and likely lies beyond the reasonable reach of current theoretical and numerical tools. Perhaps surprisingly given their considerably larger sizes, elongated DNA origamis, whose cholesteric behaviour has been recently explored experimentally<sup>100</sup>, may present a more tractable target for the application of the FY DFT due to their much greater relative stiffness. Although the persistence length of the studied 6-helix-bundle origamis is not yet as well-characterised as that of single duplexes, preliminary measurements have indeed suggested that  $l_p \geq 5l_c^{101,102}$ . Furthermore, their aspect ratio of about  $70^{100}$  leads to the formation of a stable cholesteric phase in a regime approaching the Onsager limit of high dilutions<sup>103</sup>.

We thus expect such origamis to provide a much more suitable model system in regard to the fundamental assumptions of the FY DFT. The results of their study will be presented elsewhere.

## ACKNOWLEDGMENTS

The authors are deeply grateful to A. Milchev for kindly sharing his simulation data. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 641839. The authors would like to acknowledge the use of the University of Oxford Advanced Research Computing (ARC) facility in carrying out this work. <http://dx.doi.org/10.5281/zenodo.22558>. We are grateful to the UK Materials and Molecular Modelling Hub for computational resources, which is partially funded by EPSRC (EP/P020194/1).

<sup>1</sup>R. Goetz and R. Lipowsky, J. Chem. Phys. **108**, 7397 (1998).

<sup>2</sup>C. M. Coppin and P. C. Leavis, Biophys. J. **63**, 794 (1992).

<sup>3</sup>J. W. Harris, P. Soc. Exp. Biol. Med. **75**, 197 (1950).

<sup>4</sup>D. Prevorsek, in *Polymer Liquid Crystals*, edited by A. Ciferri, W. R. Krigbaum, and R. B. Meyer (Academic Press, New York, NY, 1982) 1st ed., pp. 329 – 376.

<sup>5</sup>J.-F. Janson, “Applications of liquid-crystal polymer materials,” in *Liquid Crystal Polymers: From Structures to Applications*, edited by A. A. Collyer (Elsevier, Amsterdam, 1992) pp. 447–463.

<sup>6</sup>A. M. Donald, A. H. Windle, and S. Hanna, *Liquid Crystalline Polymers*, 2nd ed. (Cambridge University Press, Cambridge, 2006).

<sup>7</sup>A. M. Figueiredo Neto and S. R. A. Salinas, *The Physics of Lyotropic Liquid Crystals* (Oxford University Press, Oxford, 2005).

<sup>8</sup>J. P. Lagerwall and G. Scalia, Curr. Appl. Phys. **12**, 1387 (2012).

<sup>9</sup>L. Onsager, Ann. N.Y. Acad. Sci. **51**, 627 (1949).

<sup>10</sup>A. R. Khokhlov, Phys. Lett. A **68**, 135 (1978).

<sup>11</sup>A. R. Khokhlov and A. N. Semenov, Physica A **108**, 546 (1981).

<sup>12</sup>A. R. Khokhlov and A. N. Semenov, Physica A **112**, 605 (1982).

<sup>13</sup>A. R. Khokhlov and A. N. Semenov, Macromolecules **17**, 2678 (1984).

<sup>14</sup>T. Odijk, Macromolecules **19**, 2313 (1986).

<sup>15</sup>R. Hentschke, Macromolecules **23**, 1192 (1990).

<sup>16</sup>T. Sato and A. Teramoto, Mol. Cryst. Liq. Cryst. **178**, 143 (1990).

<sup>17</sup>D. B. DuPré and S.-J. Yang, J. Chem. Phys. **94**, 7466 (1991).

<sup>18</sup>Z. Y. Chen, Macromolecules **26**, 3419 (1993).

<sup>19</sup>P. van der Schoot and M. E. Cates, Europhys. Lett. **25**, 515 (1994).

<sup>20</sup>H. Fynewever and A. Yethiraj, J. Chem. Phys. **108**, 1636 (1998).

<sup>21</sup>K. M. Jaffer, S. B. Opps, D. E. Sullivan, B. G. Nickel, and L. Mederos, J. Chem. Phys. **114**, 3314 (2001).

<sup>22</sup>R. Diplock, D. E. Sullivan, K. M. Jaffer, and S. B. Opps, Phys. Rev. E **69**, 062701 (2004).

<sup>23</sup>M. Dennison, M. Dijkstra, and R. van Roij, Phys. Rev. Lett. **106**, 208302 (2011).

<sup>24</sup>W. Zhang, E. D. Gomez, and S. T. Milner, Macromolecules **48**, 1454 (2015).

<sup>25</sup>S. A. Egorov, A. Milchev, P. Virnau, and K. Binder, Soft Matter **12**, 4944 (2016).

<sup>26</sup>P.-G. de Gennes and J. Prost, *The Physics of Liquid Crystals* (Clarendon Press, Oxford, 1993).

<sup>27</sup>M. Mitov, Soft Matter **13**, 4176 (2017).

<sup>28</sup>J. P. Straley, Phys. Rev. A **14**, 1835 (1976).

<sup>29</sup>J. P. Straley, Phys. Rev. A **8**, 2181 (1973).

<sup>30</sup>M. A. Osipov, “Molecular theory of cholesteric polymers,” in *Liquid Crystalline and Mesomorphic Polymers*, edited by V. P. Shibaev and L. Lam (Springer New York, New York, NY, 1994) pp. 1–25.

<sup>31</sup>T. Odijk, J. Phys. Chem. **91**, 6060 (1987).

<sup>32</sup>R. A. Pelcovits, Liq. Cryst. **21**, 361 (1996).

<sup>33</sup>T. Sato, J. Nakamura, A. Teramoto, and M. M. Green, Macromolecules **31**, 1398 (1998).

<sup>34</sup>M. A. Osipov, Chem. Phys. **96**, 259 (1985).

<sup>35</sup>S. Varga and G. Jackson, Mol. Phys. **104**, 3681 (2006).

<sup>36</sup>H. H. Wensink and G. Jackson, J. Chem. Phys. **130**, 234911 (2009).

<sup>37</sup>H. H. Wensink and G. Jackson, J. Phys.: Condens. Matter **23**, 194107 (2011).

<sup>38</sup>H. H. Wensink, Europhys. Lett. **107**, 36001 (2014).

<sup>39</sup>M. M. C. Tortora and J. P. K. Doye, J. Chem. Phys. **146**, 184504 (2017).

<sup>40</sup>M. M. C. Tortora and J. P. K. Doye, J. Chem. Phys. **147**, 224504 (2017).

<sup>41</sup>D. Chandler, J. D. McCoy, and S. J. Singer, J. Chem. Phys. **85**, 5971 (1986).

<sup>42</sup>D. Chandler, J. D. McCoy, and S. J. Singer, J. Chem. Phys. **85**, 5977 (1986).

<sup>43</sup>J. E. Mayer and M. G. Mayer, *Statistical Mechanics* (John Wiley & Sons, Ltd, New York, NY, 1940).

<sup>44</sup>P. J. Flory, Proc. R. Soc. London A **234**, 60 (1956).

<sup>45</sup>P. J. Flory, Proc. R. Soc. London A **234**, 73 (1956).

<sup>46</sup>A. Yethiraj, Adv. Chem. Phys. **121**, 89 (2002).

<sup>47</sup>P. J. Flory, Macromolecules **11**, 1141 (1978).

- <sup>48</sup>P. J. Flory, “Molecular theory of liquid crystals,” in *Liquid Crystalline Polymers I*, edited by N. A. Platé (Springer Berlin Heidelberg, Berlin, Heidelberg, 1984) pp. 1–36.
- <sup>49</sup>G. J. Vroege and H. N. W. Lekkerkerker, Rep. Prog. Phys. **55**, 1241 (1992).
- <sup>50</sup>A. R. Khokhlov and A. N. Semenov, J. Stat. Phys. **38**, 161 (1985).
- <sup>51</sup>T. Sato and A. Teramoto, “Concentrated solutions of liquid-crystalline polymers,” in *Biopolymers Liquid Crystalline Polymers Phase Emulsion* (Springer Berlin Heidelberg, Berlin, Heidelberg, 1996) pp. 85–161.
- <sup>52</sup>J. D. Parsons, Phys. Rev. A **19**, 1225 (1979).
- <sup>53</sup>S.-D. Lee, J. Chem. Phys. **87**, 4972 (1987).
- <sup>54</sup>T. Jiang and J. Wu, J. Chem. Phys. **127**, 034902 (2007).
- <sup>55</sup>T. van Westen, T. J. Vlucht, and J. Gross, J. Chem. Phys. **137**, 044906 (2012).
- <sup>56</sup>T. van Westen, B. Oyarzún, T. J. Vlucht, and J. Gross, J. Chem. Phys. **139**, 034505 (2013).
- <sup>57</sup>J. Herzfeld, A. E. Berger, and J. W. Wingate, Macromolecules **17**, 1718 (1984).
- <sup>58</sup>M. R. Wilson, J. Mol. Liq. **68**, 23 (1996).
- <sup>59</sup>A. M. Sonnet, E. G. Virga, and G. E. Durand, Phys. Rev. E **67**, 061701 (2003).
- <sup>60</sup>A. V. Komolkin, A. Laaksonen, and A. Maliniak, J. Chem. Phys. **101**, 4103 (1994).
- <sup>61</sup>K. Šolc, J. Chem. Phys. **55**, 335 (1971).
- <sup>62</sup>D. N. Theodorou and U. W. Suter, Macromolecules **18**, 1206 (1985).
- <sup>63</sup>G. S. Grest and K. Kremer, Phys. Rev. A **33**, 3628 (1986).
- <sup>64</sup>K. Kremer and G. S. Grest, J. Chem. Phys. **92**, 5057 (1990).
- <sup>65</sup>O. Kratky and G. Porod, J. Colloid Sci. **4**, 35 (1949).
- <sup>66</sup>O. Kratky and G. Porod, Recl. Trav. Chim. Pays-Bas **68**, 1106 (1949).
- <sup>67</sup>K. G. Honnell, J. G. Curro, and K. S. Schweizer, Macromolecules **23**, 3496 (1990).
- <sup>68</sup>S. Varga and I. Szalai, Mol. Phys. **98**, 693 (2000).
- <sup>69</sup>S. A. Egorov, A. Milchev, and K. Binder, Phys. Rev. Lett. **116**, 187801 (2016).
- <sup>70</sup>E. Herrero-Galán, M. E. Fuentes-Perez, C. Carrasco, J. M. Valpuesta, J. L. Carrasosa, F. Moreno-Herrero, and J. R. Arias-Gonzalez, J. Am. Chem. Soc. **135**, 122 (2013).
- <sup>71</sup>J. R. Wenner, M. C. Williams, I. Rouzina, and V. A. Bloomfield, Biophys. J. **82**, 3160 (2002).
- <sup>72</sup>S. Geggier and A. Vologodskii, Proc. Natl. Acad. Sci. USA **107**, 15421 (2010).
- <sup>73</sup>D. H. Van Winkle, M. W. Davidson, W. X. Chen, and R. L. Rill, Macromolecules **23**, 4140 (1990).
- <sup>74</sup>F. Livolant and A. Leforestier, Prog. Polym. Sci. **21**, 1115 (1996).
- <sup>75</sup>M. Nakata, G. Zanchetta, B. D. Chapman, C. D. Jones, J. O. Cross, R. Pindak, T. Bellini, and N. A. Clark, Science **318**, 1276 (2007).
- <sup>76</sup>G. Zanchetta, F. Giavazzi, M. Nakata, M. Buscaglia, R. Cerbino, N. A. Clark, and T. Bellini, Proc. Natl. Acad. Sci. USA **107**, 17497 (2010).
- <sup>77</sup>J. P. Gouret, Biol. Cell. **32**, 299 (1978).
- <sup>78</sup>F. Livolant and M. F. Maestre, Biochemistry **19**, 3056 (1988).
- <sup>79</sup>N. S. Blanc, A. Senn, A. Leforestier, F. Livolant, and J. Dubochet, J. Struct. Biol. **134**, 76 (2001).
- <sup>80</sup>T. E. Strzelecka and R. L. Rill, Macromolecules **24**, 5124 (1991).
- <sup>81</sup>C. B. Stanley, H. Hong, and H. H. Strey, Biophys. J. **89**, 2552 (2005).
- <sup>82</sup>A. Stroobants, H. N. W. Lekkerkerker, and T. Odijk, Macromolecules **19**, 2232 (1986).
- <sup>83</sup>E. Frezza, A. Ferrarini, H. B. Kolli, A. Giacometti, and G. Cinacchi, Phys. Chem. Chem. Phys. **16**, 16225 (2014).
- <sup>84</sup>F. Tombolato and A. Ferrarini, J. Chem. Phys. **122**, 054908 (2005).
- <sup>85</sup>A. G. Cherstvy, J. Phys. Chem. B **142**, 12585 (2008).
- <sup>86</sup>S. Ruzicka and H. H. Wensink, Soft Matter **12**, 5205 (2016).
- <sup>87</sup>A. Kuhnhold and T. Schilling, J. Chem. Phys. **145**, 194904 (2016).
- <sup>88</sup>L. Wu and H. Sun, Soft Matter **14**, 344 (2018).
- <sup>89</sup>A. A. Kornyshev, S. Leikin, and S. V. Malinin, Eur. Phys. J. E **7**, 83 (2002).
- <sup>90</sup>T. E. Ouldridge, A. A. Louis, and J. P. K. Doye, J. Chem. Phys. **134**, 085101 (2011).
- <sup>91</sup>B. E. K. Snodin, F. Randisi, M. Mosayebi, P. Šulc, J. S. Schreck, F. Romano, T. E. Ouldridge, R. Tsukanov, E. Nir, A. A. Louis, and J. P. K. Doye, J. Chem. Phys. **142**, 234901 (2015).
- <sup>92</sup>J. P. K. Doye, T. E. Ouldridge, A. A. Louis, F. Romano, P. Šulc, C. Matek, B. E. K. Snodin, L. Rovigatti, J. S. Schreck, R. M. Harrison, and W. P. J. Smith, Phys. Chem. Chem. Phys. **15**, 20395 (2013).
- <sup>93</sup>F. Romano, D. Chakraborty, J. P. K. Doye, T. E. Ouldridge, and A. A. Louis, J. Chem. Phys. **138**, 085101 (2013).
- <sup>94</sup>C. Matek, T. E. Ouldridge, J. P. K. Doye, and A. A. Louis, Sci. Rep. **5**, 7655 (2015).
- <sup>95</sup>E. Skorupppa, M. Laleman, S. K. Nomidis, and E. Carlon, J. Chem. Phys. **146**, 214902 (2017).
- <sup>96</sup>K. Nadassy, I. Tomás-Oliveira, I. Alberts, J. Janin, and S. J. Wodak, Nucleic Acids Res. **29**, 3362 (2001).
- <sup>97</sup>S. Dussi, S. Belli, R. van Roij, and M. Dijkstra, J. Chem. Phys. **142**, 074905 (2015).
- <sup>98</sup>M. Dijkstra and D. Frenkel, Phys. Rev. E **51**, 5891 (1995).
- <sup>99</sup>R. Cortini, X. Cheng, and J. C. Smith, J. Phys.: Condens. Matter **29**, 084002 (2017).
- <sup>100</sup>M. Siavashpour, C. H. Wachauf, M. J. Zakhary, F. Praetorius, H. Dietz, and Z. Dogic, Nat. Mater. **16**, 849 (2017).
- <sup>101</sup>D. J. Kauert, T. Kurth, T. Liedl, and R. Seidel, Nano Lett. **11**, 5558 (2011).
- <sup>102</sup>D. Schiffels, T. Liedl, and D. K. Fyngenson, ACS Nano **7**, 6700 (2013).
- <sup>103</sup>J. P. Straley, Mol. Cryst. Liq. Cryst. **24**, 7 (1973).
- <sup>104</sup>I. Teraoka, “Models of polymer chains,” in *Polymer Solutions* (John Wiley & Sons, Inc., Hoboken, NJ, 2002) pp. 1–67.

## Appendix A: Molecular (FY) and bond (KS) order parameters

The formal definition of the bond nematic order parameter for a discretised KG chain reads as

$$S_b \equiv \left\langle \frac{3(\mathbf{t}_i \cdot \mathbf{n})^2 - 1}{2} \right\rangle_{\{\psi_{b,i}\}} \equiv \left\langle \frac{3 \cos^2 \theta_{i,i} - 1}{2} \right\rangle_{\{\psi_{b,i}\}}, \quad (\text{A1})$$

where we used the notation of Fig. 1, denoting by the brackets an average over all bond orientations  $\mathbf{t}_i$  about the uniform director  $\mathbf{n}$ . In the general case, such orientations may be described by a contour-dependent set of bond ODFs  $\{\psi_{b,i}\}$  for all bonds  $i$  comprising the chain. Neglecting the influence of the surrounding nematic field on the conformations of individual particles, one may write<sup>55,56</sup>

$$\psi_{b,i}(\mathbf{t}_i) = \psi(\mathcal{R}) \times P_{b,i}(\mathcal{R}^T \cdot \mathbf{t}_i), \quad (\text{A2})$$

with  $\psi$  the chain-averaged ODF quantifying the probability of finding a chain with orientation  $\mathcal{R}$  as defined in Fig. 1, and  $P_{b,i}$  the probability of finding the  $i$ -th bond vector pointing along  $\mathbf{t}_i$  in the molecular frame  $\mathcal{R}$  of an arbitrary chain. Under these assumptions, one thus obtains

$$\langle \cdot \rangle_{\{\psi_{b,i}\}} = \langle \langle \cdot \rangle_{\psi} \rangle_{\{P_{b,i}\}}, \quad (\text{A3})$$

where the inner brackets denote a thermodynamic average over all molecular orientations, and the outer bracket an ensemble average over all accessible bond orientations  $\mathbf{t}_i$  as expressed in the molecular frame  $\mathcal{R}$ . These two averages may therefore be computed independently by

projecting the normalised bond vectors  $\mathbf{t}_i$  onto  $\mathcal{R}$ ,

$$\begin{aligned}\mathcal{R}^\top \cdot \mathbf{t}_i &= \mathcal{R}^\top \cdot \begin{bmatrix} \sin \theta_{l,i} \cos \phi_{l,i} \\ \sin \theta_{l,i} \sin \phi_{l,i} \\ \cos \theta_{l,i} \end{bmatrix}_{\mathcal{R}_{\text{lab}}} \\ &\equiv \begin{bmatrix} \sin \theta_{m,i} \cos \phi_{m,i} \\ \sin \theta_{m,i} \sin \phi_{m,i} \\ \cos \theta_{m,i} \end{bmatrix}_{\mathcal{R}},\end{aligned}\quad (\text{A4})$$

with  $\phi_{l,i}$ ,  $\phi_{m,i}$  the azimuthal bond angles in the laboratory and molecular frames, respectively. Thus,

$$\begin{bmatrix} \sin \theta_{l,i} \cos \phi_{l,i} \\ \sin \theta_{l,i} \sin \phi_{l,i} \\ \cos \theta_{l,i} \end{bmatrix}_{\mathcal{R}_{\text{lab}}} = \mathcal{R} \cdot \begin{bmatrix} \sin \theta_{m,i} \cos \phi_{m,i} \\ \sin \theta_{m,i} \sin \phi_{m,i} \\ \cos \theta_{m,i} \end{bmatrix}_{\mathcal{R}}, \quad (\text{A5})$$

with  $\mathcal{R}$  being represented in  $\mathcal{R}_{\text{lab}}$  by the Euler angle triad  $(\alpha, \theta, \phi)$  in the  $z$ - $y$ - $z$  convention,

$$\mathcal{R} = \begin{bmatrix} \cos \alpha \cos \theta \cos \phi - \sin \alpha \sin \phi & -\sin \alpha \cos \theta \cos \phi - \cos \alpha \sin \phi & \sin \theta \cos \phi \\ \cos \alpha \cos \theta \sin \phi + \sin \alpha \cos \phi & -\sin \alpha \cos \theta \sin \phi + \cos \alpha \cos \phi & \sin \theta \sin \phi \\ -\cos \alpha \sin \theta & \sin \alpha \sin \theta & \cos \theta \end{bmatrix}_{\mathcal{R}_{\text{lab}}}. \quad (\text{A6})$$

Plugging Eq. (A6) into Eq. (A5) and projecting the result onto  $\mathbf{e}_z$  leads to

$$\begin{aligned}\cos \theta_{l,i} &= \cos \theta_{m,i} \cos \theta + \sin \theta_{m,i} \sin \theta \\ &\times (\sin \phi_{m,i} \sin \alpha - \cos \phi_{m,i} \cos \alpha),\end{aligned}\quad (\text{A7})$$

which yields, after some rearrangements,

$$\begin{aligned}\cos^2 \theta_{l,i} &= \cos^2 \theta_{m,i} \cos^2 \theta \\ &+ \sin^2 \theta_{m,i} \sin^2 \phi_{m,i} \times \sin^2 \theta \sin^2 \alpha \\ &+ \sin^2 \theta_{m,i} \cos^2 \phi_{m,i} \times \sin^2 \theta \cos^2 \alpha \\ &+ \frac{\sin 2\theta_{m,i} \sin 2\theta}{2} (\sin \phi_{m,i} \sin \alpha - \cos \phi_{m,i} \cos \alpha) \\ &- \sin 2\phi_{m,i} \sin^2 \theta_{m,i} \times \sin^2 \theta \sin \alpha \cos \alpha.\end{aligned}\quad (\text{A8})$$

Let us first perform the average of Eq. (A8) over the molecular ODF  $\psi$ , which respect to which the angles  $\theta_{m,i}$ ,  $\phi_{m,i}$  are taken to be invariant following the previous discussion. One must therefore have, for any functions

$f(\theta_{m,i}, \phi_{m,i})$  and  $g(\mathcal{R})$ ,

$$\begin{aligned}\langle f(\theta_{m,i}, \phi_{m,i}) g(\mathcal{R}) \rangle_\psi &= f(\theta_{m,i}, \phi_{m,i}) \langle g(\mathcal{R}) \rangle_\psi \\ &\equiv f(\theta_{m,i}, \phi_{m,i}) \oint d\mathcal{R} \times \psi(\mathcal{R}) g(\mathcal{R}).\end{aligned}\quad (\text{A9})$$

Furthermore, assuming the molecular ODF  $\psi$  to be cylindrically-symmetric about  $\mathbf{n}$  yields

$$\psi(\mathcal{R}) = \psi(\cos \theta), \quad (\text{A10})$$

so that for any function  $h(\theta)$ ,

$$\begin{aligned}\langle h(\theta) \rangle_\psi &= \int_0^{2\pi} d\alpha \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin \theta \times \psi(\mathcal{R}) h(\theta) \\ &= 4\pi^2 \int_0^\pi d\theta \sin \theta \times \psi(\cos \theta) h(\theta).\end{aligned}\quad (\text{A11})$$

Using Eqs. (A9) and (A11), one obtains

$$\begin{aligned}\langle \sin^2 \theta_{m,i} \sin^2 \phi_{m,i} \times \sin^2 \theta \sin^2 \alpha \rangle_\psi &= \sin^2 \theta_{m,i} \sin^2 \phi_{m,i} \int_0^{2\pi} d\alpha \times \sin^2 \alpha \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin \theta \times \psi(\cos \theta) \sin^2 \theta \\ &= \frac{\sin^2 \theta_{m,i} \sin^2 \phi_{m,i} \langle \sin^2 \theta \rangle_\psi}{2},\end{aligned}\quad (\text{A12})$$

$$\langle \sin^2 \theta_{m,i} \cos^2 \phi_{m,i} \times \sin^2 \theta \cos^2 \alpha \rangle_\psi = \frac{\sin^2 \theta_{m,i} \cos^2 \phi_{m,i} \langle \sin^2 \theta \rangle_\psi}{2}, \quad (\text{A13})$$

and it is easy to show that all the terms in the last two lines of Eq. (A8), involving only odd powers of  $\cos \alpha$  and  $\sin \alpha$ , average out to zero. Thus,

$$\langle \cos^2 \theta_{l,i} \rangle_\psi = \cos^2 \theta_{m,i} \langle \cos^2 \theta \rangle_\psi + \frac{\sin^2 \theta_{m,i} \langle \sin^2 \theta \rangle_\psi}{2}. \quad (\text{A14})$$

Finally, using the shorthand

$$\langle \cdot \rangle_P \equiv \langle \cdot \rangle_{\{P_{b,i}\}} \quad (\text{A15})$$

for the ensemble average, the bond order parameter  $S_b$  may be written as

$$\begin{aligned} S_b &= \frac{3 \langle \cos^2 \theta_{l,i} \rangle_\psi - 1}{2} \\ &= \frac{3 \langle \cos^2 \theta_{m,i} \rangle_P \langle \cos^2 \theta \rangle_\psi}{2} + \frac{3 \langle \sin^2 \theta_{m,i} \rangle_P \langle \sin^2 \theta \rangle_\psi}{4} - \frac{1}{2} \\ &= \frac{3 \langle \cos^2 \theta_{m,i} \rangle_P \langle \cos^2 \theta \rangle_\psi}{2} + \frac{3 (\langle \sin^2 \theta_{m,i} \rangle_P - 1) (\langle \sin^2 \theta \rangle_\psi - 1)}{4} + \frac{3 (\langle \sin^2 \theta_{m,i} \rangle_P + \langle \sin^2 \theta \rangle_\psi)}{4} - \frac{5}{4} \\ &= \frac{9 \langle \cos^2 \theta_{m,i} \rangle_P \langle \cos^2 \theta \rangle_\psi}{4} - \frac{3 (\langle \cos^2 \theta_{m,i} \rangle_P + \langle \cos^2 \theta \rangle_\psi)}{4} + \frac{1}{4} \\ &= \left( \frac{3 \langle \cos^2 \theta_{m,i} \rangle_P - 1}{2} \right) \times \left( \frac{3 \langle \cos^2 \theta \rangle_\psi - 1}{2} \right). \end{aligned} \quad (\text{A16})$$

Thus,

$$S_b = S \times S_{\text{int}}, \quad (\text{A17})$$

with  $S$  the standard nematic order parameter for the molecular long axes,

$$S = \frac{3 \langle \cos^2 \theta \rangle_\psi - 1}{2}, \quad (\text{A18})$$

and  $S_{\text{int}}$  an ensemble-averaged order parameter quantifying the distribution of bond orientations about the long axis of the chain conformations,

$$S_{\text{int}} = \frac{3 \langle \cos^2 \theta_{m,i} \rangle_P - 1}{2}, \quad (\text{A19})$$

which yield the discretised versions of Eqs. (14)–(17).

## Appendix B: Bending stiffness and intra-molecular order parameter for unconfined chains

Let us now work out the scaling behaviour of the order parameter  $S_{\text{int}}$  with chain contour length  $l_c$  and persistence length  $l_p$  in the case of unconfined chains. In the KG bead-spring model, chain stiffness is governed by the bond bending potential,

$$U_{\text{bend}} = \sum_{i=1}^{N-2} \epsilon_b (1 - \mathbf{t}_{i+1} \cdot \mathbf{t}_i) = \frac{\epsilon_b}{2} \sum_{i=1}^{N-2} (\mathbf{t}_{i+1} - \mathbf{t}_i)^2, \quad (\text{B1})$$

with  $\mathbf{t}_i$  the normalised bond vector linking the monomers  $i$  and  $i+1$ , respectively located in  $\mathbf{r}_i$  and  $\mathbf{r}_{i+1}$ ,

$$\mathbf{t}_i \equiv \frac{\mathbf{r}_{i+1} - \mathbf{r}_i}{\|\mathbf{r}_{i+1} - \mathbf{r}_i\|}, \quad (\text{B2})$$

where  $\|\cdot\|$  is the Euclidean norm. Let us denote by  $l_b$  the average bond length separating adjacent monomers, taken to be a constant, and define  $s \equiv l_b \times i$  as the curvilinear abscissa of the  $i$ -th monomer. In the continuum limit  $l_c \gg l_b$ , one may assume  $\{\mathbf{r}(s), \mathbf{t}(s)\} \equiv \{\mathbf{r}_i, \mathbf{t}_i\}$  to be differentiable functions of the continuous variable  $s$ , so that Eq. (B1) converges towards the Riemann integral

$$U_{\text{bend}} = \frac{\epsilon_b l_b}{2} \int_0^{l_c} ds \times \left( \frac{d\mathbf{t}}{ds} \right)^2, \quad (\text{B3})$$

and Eq. (B2) takes the form of the simple differential

$$\mathbf{t} = \frac{d\mathbf{r}}{ds}, \quad (\text{B4})$$

with  $\mathbf{r}(s)$  the continuous curve describing the chain conformation in space. One thus recovers the standard worm-like chain model, with bending stiffness  $K = \epsilon_b l_b \equiv l_p k_b T$ .

Let us denote by  $\mathcal{R} \equiv [\mathbf{v} \times \mathbf{u} \quad \mathbf{v} \quad \mathbf{u}]$  the three unit vectors defining the molecular frame  $\mathcal{R}$ . In the framework of PCA,  $\mathbf{v}$  and  $\mathbf{u}$  respectively correspond to the molecular short and long axes, and are defined such that the following inequalities are always verified,

$$\lambda_{\mathbf{u}}^2 > \lambda_{\mathbf{v} \times \mathbf{u}}^2 > \lambda_{\mathbf{v}}^2, \quad (\text{B5})$$



where the  $\lambda_{\mathbf{x}}^2$  represent the principal moments of the gyration tensor, and quantify the extent of the chain along axis  $\mathbf{x}$ . Their expression reads, in our notation,

$$\lambda_{\mathbf{x}}^2 \equiv \langle (\mathbf{r} \cdot \mathbf{x})^2 \rangle_c \quad \forall \mathbf{x} \in \{\mathbf{v} \times \mathbf{u}, \mathbf{v}, \mathbf{u}\}, \quad (\text{B6})$$

with  $\langle \cdot \rangle_c$  the contour average,

$$\langle \cdot \rangle_c \equiv \frac{1}{l_c} \int_0^{l_c} ds \cdot, \quad (\text{B7})$$

assuming the particle centre of mass to be set to the origin of the frame, i.e.,  $\langle \mathbf{r} \rangle_c = \mathbf{0}$ . The adimensioned molecular anisotropy parameter  $\kappa$  is then defined as<sup>62</sup>

$$\kappa \equiv \frac{3}{2} \frac{\lambda_{\mathbf{u}}^4 + \lambda_{\mathbf{v}}^4 + \lambda_{\mathbf{v} \times \mathbf{u}}^4}{(\lambda_{\mathbf{u}}^2 + \lambda_{\mathbf{v}}^2 + \lambda_{\mathbf{v} \times \mathbf{u}}^2)^2} - \frac{1}{2}, \quad (\text{B8})$$

which verifies  $0 \leq \kappa \leq 1$ ,  $\kappa = 0$  being reached in the limit of a spherically-symmetric monomer distribution and  $\kappa = 1$  in the case of an ideal linear chain.

In the following, we restrict our study to systems of stiff polymers such that  $l_p \gg l_c$ , for which  $\kappa \sim 1$ . In this case, the first inequality in Eq. (B5) becomes wide, so that

$$\lambda_{\mathbf{u}}^2 \gg \lambda_{\mathbf{v} \times \mathbf{u}}^2, \lambda_{\mathbf{v}}^2. \quad (\text{B9})$$

One may then approximate the end-to-end separation vector by its projection onto  $\mathbf{u}$ ,

$$[\mathbf{r}(l_c) - \mathbf{r}(0)] \cdot \mathbf{u} = \int_0^{l_c} ds \times \mathbf{t}(s) \cdot \mathbf{u} = \int_0^{l_c} ds \times \cos \theta_m(s), \quad (\text{B10})$$

where we used Eq.(B4), denoting by  $\theta_m(s)$  the continuous limit of the bond angle  $\theta_{m,i}$  as defined in Fig. 1. The mean square end-to-end separation distance  $\langle L^2 \rangle$  thus reads as

$$\langle L^2 \rangle \simeq \int_0^{l_c} ds \int_0^{l_c} ds' \times \langle \cos \theta_m(s) \cos \theta_m(s') \rangle. \quad (\text{B11})$$

Neglecting the contour fluctuations of  $\theta_m(s)$  for stiff polymers, one may rather crudely write

$$\cos \theta_m(s) \simeq \cos \theta_m(s') \simeq \langle \cos \theta_m \rangle_c, \quad (\text{B12})$$

so that Eq. (B11) yields

$$\begin{aligned} \langle L^2 \rangle &\simeq \int_0^{l_c} ds \int_0^{l_c} ds' \times \langle \langle \cos \theta_m \rangle_c^2 \rangle \\ &\simeq l_c^2 \times \langle \langle \cos^2 \theta_m \rangle_c \rangle. \end{aligned} \quad (\text{B13})$$

Recognising in Eq. (B13) the continuum equivalent of the ensemble average  $\langle \cdot \rangle_P$  as defined in Eq. (A15), one may recast Eq. (A19) in the form

$$S_{\text{int}} = \frac{3 \langle \langle \cos^2 \theta_m \rangle_c \rangle - 1}{2} \simeq \frac{3 \langle L^2 \rangle}{2 l_c^2} - \frac{1}{2}. \quad (\text{B14})$$

For inextensible and unconfined worm-like chains as described by Eqs. (B3) and (B4), the algebraic expression for  $\langle L^2 \rangle$  reads as<sup>104</sup>

$$\frac{\langle L^2 \rangle}{l_c^2} = 2 \left( \frac{l_p}{l_c} \right)^2 \times \left[ \frac{l_c}{l_p} - 1 + \exp \left( - \frac{l_c}{l_p} \right) \right] \quad (\text{B15})$$

$$= 1 - \frac{l_c}{3l_p} + \mathcal{O} \left( \left( \frac{l_c}{l_p} \right)^2 \right), \quad (\text{B16})$$

which yields a simple scaling law for the bond order parameter of highly-stiff chains,

$$S_{\text{int}} \simeq 1 - \frac{l_c}{2l_p}. \quad (\text{B17})$$

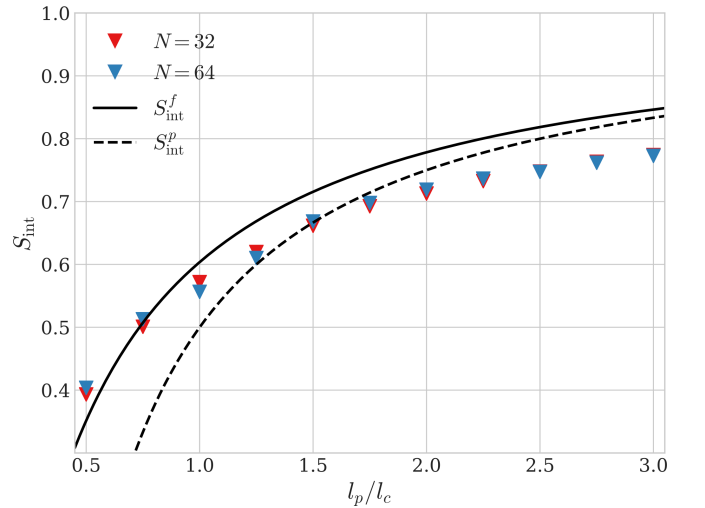


FIG. 5. Intra-molecular bond order parameter  $S_{\text{int}}$  as a function of persistence-to-contour-length ratio for unconfined semi-flexible bead-spring chains. Markers represent the results of Metropolis MC simulations for chains with  $N = 32$  and  $N = 64$  monomers following the procedure of Sec. III.  $S_{\text{int}}^f$  denotes the theoretical predictions obtained by plugging the full Eq. (B15) into Eq. (B14), and  $S_{\text{int}}^p$  those of the partial asymptotic expansion Eq. (B17).