

## Self-Assembly of Bifunctional Patchy Particles with Anisotropic Shape into Polymers Chains: Theory, Simulations, and Experiments

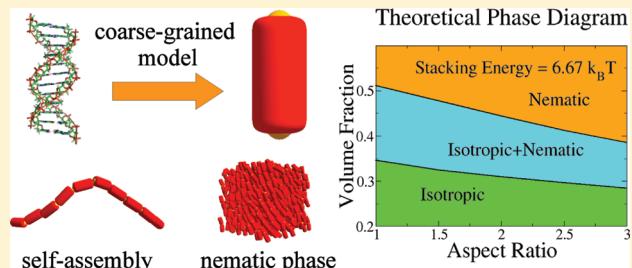
Cristiano De Michele,<sup>\*,†</sup> Tommaso Bellini,<sup>‡</sup> and Francesco Sciortino<sup>§</sup>

<sup>†</sup>Dipartimento di Fisica, Sapienza - Università di Roma, P. le A. Moro 2, 00185 Roma, Italy

<sup>‡</sup>Dipartimento di Chimica, Biochimica e Biotecnologie per la Medicina, Università di Milano, I-20122 Milano, Italy

<sup>§</sup>Dipartimento di Fisica and CNR-ISC, Sapienza - Università di Roma, P. le A. Moro 2, 00185 Roma, Italy

**ABSTRACT:** Concentrated solutions of short blunt-ended DNA duplexes, as short as 6 base pairs, are known to order into the nematic liquid crystal phases. This self-assembly is due to the stacking interactions between duplex terminals that promotes their aggregation into polydisperse chains with a significant persistence length. Experiments show that liquid crystal phases form above a critical volume fraction depending on the duplex length. We introduce and investigate via numerical simulations, a coarse-grained model of DNA double-helical duplexes. Each duplex is represented as an hard quasi-cylinder whose bases are decorated with two identical reactive sites. The stacking interaction between terminal sites is modeled via a short-range square-well potential. We compare the numerical results with predictions based on a free energy functional and find satisfactory quantitative matching of the isotropic–nematic phase boundary and of the system structure. Comparison of numerical and theoretical results with experimental findings confirm that the DNA duplex self-assembly can be properly modeled via equilibrium polymerization of cylindrical particles. This insight enables us to estimate the stacking energy.



### I. INTRODUCTION

Self-assembly is the spontaneous organization of matter into reversibly bound aggregates. In contrast to chemical synthesis, where molecular complexity is achieved through covalent bonds, supramolecular aggregates spontaneously form in a self-assembly process due to free energy minimization. Self-assembly is ubiquitous in nature and can involve the structuring of elementary building blocks of various sizes, ranging from simple molecules (e.g., surfactants) to the mesoscopic units (e.g., colloidal particles). This ability to assemble complicated entities from relatively simple building blocks has made self-assembly interesting to several fields, including soft matter and biophysics.<sup>1–3</sup> Understanding and thus controlling self-assembly is important for devising new materials whose physical properties are controlled by tuning the interactions of the various components.<sup>4–7,7–12</sup>

A particular but very interesting case of self-assembly occurs when the anisotropy of attractive interactions between the monomers favors the formation of linear or filamentous aggregates, i.e., linear chains. A longstanding example is provided by the formation of worm-like micelles of amphiphilic molecules in water or microemulsions of water and oil which are stabilized by amphiphilic molecules. If supramolecular aggregates possess a sufficient rigidity the system may exhibit liquid crystal (LC) ordering even if the self-assembling components do not have the required shape anisotropy to guarantee the formation of nematic phases. Intense experimental activity has been dedicated to the study of nematic transitions in micellar systems.<sup>13–15</sup> Another prominent case is

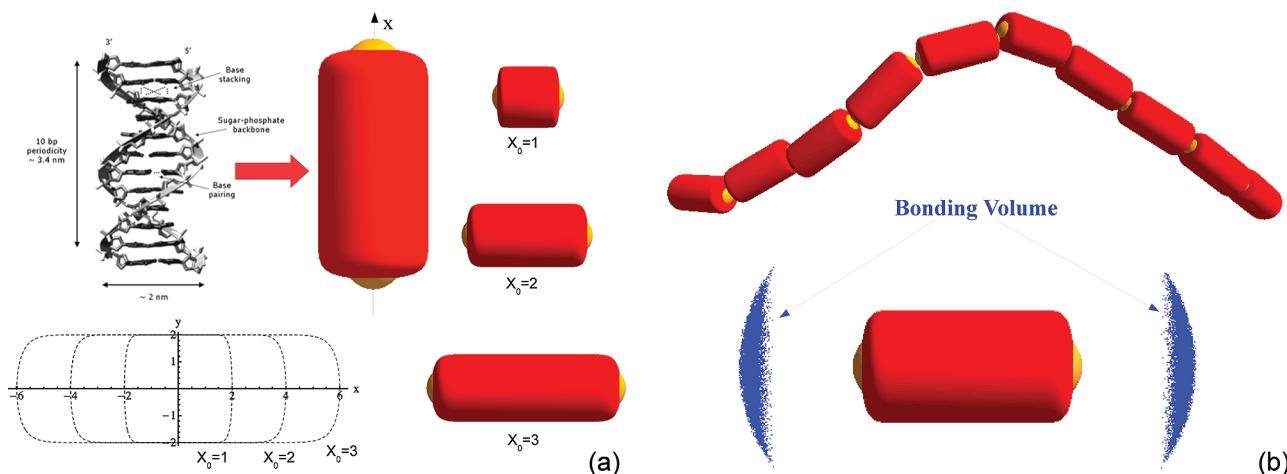
the formation of fibers and fibrils of peptides and proteins.<sup>16–19</sup> Over the last 50 years LC phases have been also observed in solutions of long duplex B-form DNA composed of  $10^2$  to  $10^6$  base pairs,<sup>20–23</sup> and in the analogous case of filamentous viruses.<sup>24–28</sup> More recently, a series of experiments<sup>29–31</sup> have provided evidence that a solution of short DNA duplexes (DNAD), 6 to 20 base pairs in length, can also form liquid crystals above a critical concentration, giving rise to nematic and columnar LC phases.<sup>29</sup>

However, this behavior was found when the terminals of the duplexes interact attractively. This condition is verified either when duplexes terminate bluntly, as in the case of fully complementary strands shown in Figure 1a, or when the strands arrange in shifted double-helices whose overhangs are mutually interacting. This behavior is not restricted to B-form DNA oligomers, as it has also been observed in solutions of blunt-ended A-form RNA oligomeric duplexes.<sup>32</sup> As the terminal groups are modified to disrupt attraction, the LC long-range ordering is lost. Overall, the whole body of experimental evidence supports the notion that LC formation is due to the formation of reversible linear aggregates of duplexes, in turn promoting the onset of long-ranged orientational LC ordering. According to this picture, the LC ordering of oligomeric DNA is analogous to the LC ordering of chromonic liquid crystals.<sup>33</sup> Both in chromonics and in blunt-

Received: August 27, 2011

Revised: December 12, 2011

Published: December 29, 2011



**Figure 1.** Coarse-grained model of DNA duplexes. (a) DNA duplex and a 3D graphical representation of its corresponding coarse-grained model comprising a SQ<sub>1</sub> symmetric around the  $x$  axis, decorated with two sticky spots located on its bases. The figure also shows SQs of different aspect ratios ( $X_0 = 1, 2, 3$ ) and the projection of their surfaces onto the  $xy$ -plane. Note that the base roundness increases on increasing  $X_0$ . (b) Random chain of 10 monomers and a representation (blue clouds) of the points where the center of mass of a different monomer can be located in a bonding configuration. This set of points defines the bonding volume.

ended DNA duplexes, the aggregation takes place because of stacking interactions, generally understood as hydrophobic forces acting between the flat hydrocarbon surfaces provided by the core of chromonic molecules and by the paired nucleobases at the duplex terminals.<sup>34,35</sup>

The LC ordering of nucleic acids is relevant for various reasons. First, it provides a new model of reversible aggregation leading to macroscopic ordering in which the strength of the intermonomer attraction can be modified by changing the duplex terminals (blunt-end stacking or pairing of overhangs). Second, it provides new access to DNA–DNA interactions, and in particular to stacking interactions, whose nature is still being investigated and debated.<sup>34,35</sup> In this vein, self-assembly acts as an amplifier of the intermonomeric interactions, enabling study of the effects of minor molecular modification (e.g., oligomer terminations) on base stacking. Finally, stacking and self-assembly are often invoked as the prebiotic route to explaining the gap between the random synthesis of elementary carbon-based molecules and the first complex molecules, possibly RNA oligomers, which are capable of catalyzing their own synthesis.<sup>36</sup> To proceed in any of these directions, it is necessary to rely on models which allow us to quantitatively connect the collective behavior of nucleic acids oligomers to their molecular properties and, in particular, to the duplex size and to the strength and range of the interduplex attractions.

While the isotropic–nematic transition in rigid and semi-flexible polymers has been investigated in details in the past and rather accurate thermodynamic descriptions have been proposed,<sup>37–45</sup> much less is known for the case in which the nematic transition takes place in reversibly assembling polymers, i.e., when the average length of the chains depends on the state point explored. Recent theoretical and numerical works<sup>46,47</sup> has renewed the interest in this topic.<sup>48</sup> Reference 47 investigates the self-assembly and nematization of spheres, while ref 46 focuses on the polymerization of interacting cylinders. In this article, we propose a coarse-grained model similar to the one introduced in ref 46 devised to capture the essential physical features of equilibrium polymerization of DNA duplexes. We study this model numerically via Monte Carlo simulations in the constant temperature and pressure ensembles, applying special biasing techniques<sup>49,50</sup> to speed up

the equilibration process. We then develop a free-energy functional, building on Wertheim<sup>51–53</sup> and Onsager<sup>54</sup> theories, which provides a satisfactory description of the system in the isotropic and nematic phases. A comparison of the calculated phase boundaries for different aspect ratios and different interaction strengths with experimental results allow us to confirm that (i) the DNAD aggregation and LC ordering processes can be properly modeled via equilibrium polymerization of cylindrical particles and (ii) to provide an estimate of the stacking energy.

In section II, we introduce the coarse-grained model of DNADs and provide some details of the computer simulations. Section III gives a summary of the analytic theory which we developed to describe the system in the isotropic and nematic phases. A comparison of our analytical approach with numerical results is presented in section V, while in section VI we provide an estimate of the stacking energy by comparing our theoretical results with experimental data. Section VII represents our conclusions.

## II. MODEL AND NUMERICAL DETAILS

In this section we introduce a coarse-grained model devised to capture the essential physical features of end-to-end stacking (equilibrium polymerization) of DNA duplexes which is well suited to being investigated both theoretically and numerically. In the model, particles (DNADs) are assimilated to superquadrics (SQ) with a quasi-cylindrical shape decorated with two reactive sites on their bases determining their interactions. SQs are a straightforward generalization of hard ellipsoids (HE), their surface is in fact defined as follows:

$$f(x, y, z) = \left| \frac{x}{a} \right|^p + \left| \frac{y}{b} \right|^m + \left| \frac{z}{c} \right|^n - 1 = 0 \quad (1)$$

where the parameters  $p, m, n$  are real numbers and  $a, b, c$  are the SQ semiaxes.

In our case we set  $m = n = 2$ ,  $p = 16$ , and  $b = c$ , so that the SQ resembles a cylinder with rounded edges (see Figure 1). The absence of surface gradient discontinuities makes this model also suitable for investigating its dynamics via event-driven molecular dynamics.<sup>55</sup>

Such SQs can be fully characterized by the aspect ratio  $X_0 = a/b$  and by the parameter  $p$ , that determines the sharpness of the edges (see Figure 1). As for the case of HEs SQs of aspect ratio  $X_0 < 1$  are called “oblate”, while SQs of aspect ratio  $X_0 > 1$  are called “prolate”. As unit of length in our simulations we use the length of the short semiaxes  $b$ . In the present study we investigated only prolate SQs with aspect ratio  $X_0 = 1.2$  and 3. We chose such aspect ratios because DNADs used in experiments<sup>29</sup> have a diameter  $D = 2$  nm and are composed of 6 to 20 base pairs (BP) each 0.3 nm long. Hence their aspect ratio  $X_0$  ranges approximately from 1 to 3.

Each particle is decorated with two attractive sites, located along the symmetry axis ( $x$ -axis in Figure 1) at a distance  $d/b = X_0 - 0.46$  from the DNAD center of mass, in order to model hydrophobic (stacking) forces between DNADs. Sites belonging to distinct particles interact via the following square-well (SW) potential:

$$\beta u_{SW} = \begin{cases} -\beta\Delta E_S & r < \delta \\ 0 & r > \delta \end{cases} \quad (2)$$

Here  $r$  is the distance between the interacting sites,  $\delta/b = 1.22$  is the range of interaction (i.e., the diameter of the attractive sites),  $\beta = 1/k_B T$ , and  $k_B$  is the Boltzmann constant. Therefore, in the present model the anisotropic hard-core interaction is complemented with an anisotropic attractive potential in a fashion similar to past work on water,<sup>56</sup> silica,<sup>57</sup> and the stepwise polymerization of bifunctional diglycidyl ether of bisphenol A with pentafunctional diethylenetriamine.<sup>58,59</sup>

The location and diameter of the attractive sites have been chosen to best mimic the stacking interactions between blunted-ended DNAD, and in particular, they ensure the following points.

- 1 The maximum interaction range between two DNAD bases is of the order of the typical range for hydrophobic interactions (i.e., 2 Å, see ref 60), i.e., comparable to the dimensions of the water molecule.
- 2 The extent of the attractive surface of the DNAD bases is compatible with the surface of aromatic groups present in DNADs, which are responsible for hydrophobic interactions.

We note that in the present model each DNAD is symmetric around the  $x$ -axis (see Figure 1), and hence, we are neglecting rotations around it.

We performed Monte Carlo (MC) simulations in the canonical and isobaric ensembles. We implemented the aggregation biased MC technique (AVBMC) developed by Chen and Siepmann<sup>49,50</sup> in order to speedup (up to a factor of 2 compared to standard MC) the formation of linear aggregates.

To detect the overlap of two DNADs we calculated the distance using the algorithm described in ref 55. In all simulations we adopted periodic boundary conditions in a cubic simulation box.

We studied a system of  $N = 1000$  particles in a wide range of volume fractions  $\phi$  and pressure  $P$ , respectively. Initially, we prepared configurations at high temperature with all DNADs being bonded, and then we quenched the system to the final temperature (i.e., to the final value of  $\beta\Delta E_S$ ) before letting it equilibrate. We checked equilibration by inspecting the behavior of the potential energy and the nematic order parameter (see section VB) in the system.

### III. THEORY

Following the work of van der Schoot and Cates<sup>14,48</sup> and its extension to higher volume fractions with the use of the Parsons–Lee approximation<sup>61,62</sup> as suggested by Kuriabova et al.,<sup>46</sup> we assume the following expression for the free energy of our system:

$$\frac{\beta F}{V} = \sum_{l=1}^{\infty} \nu(l) \{ \ln[\nu_d \nu(l)] - 1 \} + \frac{\eta(\phi)}{2} \sum_{l=1}^{\infty} \sum_{l'=1}^{\infty} \nu(l) \nu(l') \nu_{excl}(l, l') - (\beta\Delta E_S + \sigma_b) \sum_{l=1}^{\infty} (l-1) \nu(l) + \sum_{l=1}^{\infty} \nu(l) \sigma_o(l) \quad (3)$$

where  $\nu(l)$  is the number density of chains of length  $l$ , normalized such that  $\sum_{l=1}^{\infty} l \nu(l) = \rho$ ,  $\nu_d$  is the volume of a monomer,  $\beta\Delta E_S$  is the (positive) stacking energy,  $\nu_{excl}(l, l')$  is the excluded volume of two chains of length  $l$  and  $l'$  and  $\sigma_b$  is the entropic free energy penalty for bonding (i.e., is the contribution to free energy due to the entropy which is lost by forming a single bond).  $\eta(\phi)$  is the Parsons–Lee factor<sup>61</sup>

$$\eta(\phi) = \frac{1}{4} \frac{4 - 3\phi}{(1 - \phi)^2} \quad (4)$$

and  $\sigma_o$ <sup>45</sup> accounts for the orientational entropy that a chain of length  $l$  loses in the nematic phase (including possible contribution due to its flexibility). Different from refs 46 and 48 but as in ref 47, we explicitly account for the polydispersity inherent in the equilibrium polymerization using a discrete chain length distribution. We explicitly separate the bonding free energy in an energetic ( $\beta\Delta E_S$ ) and an entropic ( $\sigma_b$ ) contribution. Different from ref 47 and 48 but as in ref 46, we include the Parsons–Lee factor. Indeed, the Parsons decoupling approximation satisfactory models the phase diagram of uniaxial hard ellipsoids,<sup>63</sup> hard cylinders,<sup>64</sup> linear fused hard spheres chains,<sup>65</sup> mixtures of hard platelets,<sup>66</sup> hard spherocylinders,<sup>67–69</sup> rod–plate mixtures,<sup>70</sup> mixtures of rod-like particles<sup>71,72</sup> and mixtures of hard rods and hard spheres.<sup>73</sup> On the other hand, ref 74 finds that the Parsons theory is not satisfactory in the case of rigid linear chains of spheres.

A justification of the use of Parsons–Lee factor in eq 3 for the present case of aggregating cylinders is provided in Appendix A. Here we only note that the present system, in the limit of high  $T$  where polymerization is not effective, reduces to a fluid of hard quasi-cylinders, where the use of Parsons–Lee factor is justifiable.<sup>64,68,69</sup> Moreover, in the dilute limit ( $\eta(\phi) \rightarrow 1$ ) the excluded volume term in eq 3 reduces to the excluded volume of a polydisperse set of aggregates with length distribution  $\nu(l)$ , which conforms to Onsager’s original theory.<sup>54</sup> In other words, the form chosen in eq 3 for the excluded volume contribution to the free energy reduces to the correct expressions in the limit of high temperatures and of low volume fractions.

Following van der Schoot and Cates,<sup>14,48</sup>  $\nu_{excl}(l, l')$  can be assumed to be as a second order polynomial in  $l$  and  $l'$

$$\nu_{excl}[l, l'; f(\mathbf{u})] = 2 \int f(\mathbf{u})f(\mathbf{u}') D^3[\Psi_1(\gamma, X_0) + \frac{l+l'}{2}\Psi_2(\gamma, X_0)X_0 + \Psi_3(\gamma, X_0)X_0^2 ll'] d\Omega d\Omega' \quad (5)$$

where  $f(\mathbf{u})$  is the probability for a given monomer having an orientation  $\mathbf{u}$  within the solid angle  $\Omega$  and  $\Omega + d\Omega$  and  $\Psi_\alpha$  describes the angular dependence of the excluded volume. The orientational probability  $f(\mathbf{u})$  is normalized as

$$\int f(\mathbf{u}) d\Omega = 1 \quad (6)$$

In particular, for two rigid chains of length  $l$  and  $l'$  which are composed of hard cylinders (HC) of diameter  $D$  and length  $X_0D$ ,  $\nu_{excl}(l, l')$  has been calculated by Onsager in 1949

$$\begin{aligned} \nu_{excl}(l, l') = & \int f(\mathbf{u})f(\mathbf{u}') D^3 \left[ \frac{\pi}{2} \right. \\ & \sin \gamma + \frac{\pi}{2} X_0 (1 + |\cos \gamma| \\ & + \frac{4}{\pi} E(\sin \gamma)) \frac{l+l'}{2} + 2X_0^2 \sin \gamma ll' \left. \right] \\ & d\Omega d\Omega' \end{aligned} \quad (7)$$

where  $\cos \gamma = \mathbf{u} \cdot \mathbf{u}'$  and  $E(\sin \gamma)$  is the complete elliptical integral

$$E(\sin \gamma) = \frac{1}{4} \int_0^{2\pi} (1 - \sin^2 \gamma \sin^2 \psi)^{1/2} d\psi \quad (8)$$

In passing, we observe that the integrals in eq 7 can be calculated exactly in the isotropic phase while in the nematic phase the calculation can be done analytically only with suitable choices of the angular distribution  $f(\mathbf{u})$ . Comparing eqs 7 and 5 for HC one has:

$$\begin{aligned} \Psi_1(\gamma, X_0) &= \frac{\pi}{4} \sin \gamma \\ \Psi_2(\gamma, X_0) &= \frac{\pi}{4} \left( 1 + |\cos \gamma| + \frac{4}{\pi} E(\sin \gamma) \right) \\ \Psi_3(\gamma, X_0) &= \sin \gamma \end{aligned} \quad (9)$$

In view of eqs 9 we note that for HCs the functions  $\Psi_1(\gamma)$ ,  $\Psi_2(\gamma)$ , and  $\Psi_3(\gamma)$  accounts for the orientational dependence of the excluded volume of two monomers having orientations  $\mathbf{u}$  and  $\mathbf{u}'$  with  $\mathbf{u} \cdot \mathbf{u}' = \cos \gamma$ . It is also worth observing that the first term of the integrand in eq 7 is independent of  $l$  and hence accounts for the excluded volume interaction between two HCs ends. The second term is linear in  $l$  and  $l'$  and accounts for the excluded volume between the end of a chain and all midsections of a second one. The third term, proportional to  $ll'$ , models the interaction between all  $ll'$  pairs of midsections of the two chains.<sup>14,48</sup> In summary, eq 5 is exact for two rigid chains of HCs but, to lowest order of approximation,<sup>14,39</sup> it is justifiable even for two semiflexible chains. We then assume that  $\nu_{excl}$  remains additive with respect to end-end, end-midsection and midsection-midsection excluded volume contributions even if the chain is semiflexible. Finally, our further ansatz is that eq 5 is also a good functional form for the excluded volume of two superquadratics having quasi-cylindrical

shape: we will check the validity of this hypothesis using our simulations data.

It is worth noting that in the present case eq 7 is not appropriate for evaluating the excluded volume between two linear aggregates of SQs, for at least two reasons: (i) a superquadric and a cylinder with same diameter and aspect ratio have different volumes. (ii) Equation 7 holds if linear aggregates of SQs retain a “tube-like” shape. In the present model instead, two bonded SQs may have their symmetry axes parallel but not coaxial. Despite the general inadequacy of eq 7 for our SQ model, one can expect that the scaling with respect to  $l$  and  $l'$  is still valid and for this reason we assume the functional form in eq 5 for the excluded volume between two chains of SQs.

An exact expression for  $\sigma_o$  is not available. The two following limits have been calculated by Khokhlov and Semenov:<sup>37,40,45</sup>

$$\begin{aligned} \sigma_o(l) = & \frac{l}{8l_p} \int \left( \frac{\partial f}{\partial \theta} \right)^2 f^{-1} d\Omega - 2 \ln \left[ \int f^{1/2} d\Omega \right] \\ & + \ln(4\pi) \quad (l_p \ll l) \\ \sigma_o(l) = & \int f \ln(4\pi f) d\Omega + \frac{l}{12l_p} \int \left( \frac{\partial f}{\partial \theta} \right)^2 f^{-1} d\Omega \\ & (l_p \gg l) \end{aligned} \quad (10)$$

Finally, we note that, in the limit of rigid rods with  $f_l(\mathbf{u}) = f(\mathbf{u})\nu(l)$ , (the same limit selected in ref 46), the free energy in eq 3 reduces to:

$$\begin{aligned} \frac{\beta F}{V} = & \frac{2}{3l_p} \sum_{l=1}^{\infty} l \int [f_l(\mathbf{u})]^{1/2} \nabla^2 [f_l(\mathbf{u})]^{1/2} d\Omega + \\ & \sum_{l=1}^{\infty} \int f_l(\mathbf{u}) \{ \ln[4\pi\nu_d f_l(\mathbf{u})] - 1 \} \\ & + \frac{\eta(\varphi)}{2} \sum_{l=1, l'=1}^{\infty} \int f_l(\mathbf{u}) f_{l'}(\mathbf{u}') \nu_{excl}(l, l') \\ & d\Omega d\Omega' - (\beta \Delta E_S + \sigma_b) \\ & \sum_{l=1}^{\infty} \int (l-1) f_l(\mathbf{u}) d\Omega \end{aligned} \quad (11)$$

which is analogous to the free energy expression used by Kuriabova et al.<sup>46</sup>

**A. Isotropic Phase.** In the isotropic phase all orientations are equiprobable, and hence

$$f(\mathbf{u}) = \frac{1}{4\pi} \quad (12)$$

Plugging eq 12 into eq 3 and calculating the integrals one obtains:

$$\begin{aligned} \frac{\beta F}{V} = & \sum_{l=1}^{\infty} \nu(l) \{ \ln[\nu_d \nu(l)] - 1 \} + \frac{\eta(\varphi)}{2} \\ & \sum_{l=1, l'=1}^{\infty} \nu(l) \nu(l') \nu_{excl}(l, l') - (\beta \Delta E_S + \sigma_b) \\ & \sum_{l=1}^{\infty} (l-1) \nu(l) \end{aligned} \quad (13)$$

For hard cylinders the excluded volume can be calculated explicitly:

$$\nu_{excl}(l, l') = \frac{\pi^2}{8} D^3 + \left( \frac{3\pi}{8} + \frac{\pi^2}{8} \right) [l + l'] X_0 D^3 + \frac{\pi}{2} l' X_0^2 D^3 \quad (14)$$

Building on eq 14, the generic expression for the excluded volume  $\nu_{excl}(l, l')$  reported in eq 5 in the isotropic phase takes the form:

$$\nu_{excl}(l, l') = 2 \left[ A_I(X_0) + k_I(X_0) v_d \frac{l + l'}{2} + B_I(X_0) X_0^2 l' \right] \quad (15)$$

We assume that the chain length distribution  $\nu(l)$  is exponential with an average chain length  $M$

$$\nu(l) = \rho M^{-(l+1)} (M - 1)^{l-1} = \frac{\rho}{M(M-1)} e^{-l[\ln M - \ln(M-1)]} \quad (16)$$

where

$$M = \frac{\sum_l^\infty l \nu(l)}{\sum_l^\infty \nu(l)} \quad (17)$$

With this choice for  $\nu(l)$  the free energy in eq 13 becomes:

$$\begin{aligned} \frac{\beta F}{V} = & -\rho(\beta\Delta E_S + \sigma_b)(1 - M^{-1}) \\ & + \eta(\phi) \left[ B_I X_0^2 + \frac{v_d k_I}{M} + \frac{A_I}{M^2} \right] \rho^2 \\ & + \frac{\rho}{M} \left[ \ln \left( \frac{v_d \rho}{M} \right) - 1 \right] + \rho \frac{M-1}{M} \ln(M-1) - \rho \ln M \end{aligned} \quad (18)$$

Note that, in general,  $k_I$ ,  $B_I$  and  $A_I$  depend on  $X_0$ .

The minimization of the free energy with respect to  $M$  yields, after dropping terms in  $O(1/M^2)$

$$M = \frac{1}{2} (1 + \sqrt{1 + 4\omega\phi e^{k_I\phi\eta(\phi)+\beta\Delta E_S}}) \quad (19)$$

where  $\omega \equiv e^{\sigma_b}$ . This formula for  $M$  differs from the one reported by Kindt<sup>47</sup> by the presence of the Parsons-Lee factor, which will play a role at high volume fractions.

The expression for  $M$  in eq 19 coincides with the parameter-free expression for the average chain length  $M_w$  obtained within Wertheim's theory (e.g., see refs 51–53, 75, and 76), when  $\phi$  is small and  $e^{k_I\phi\eta(\phi)} \approx 1$ . Indeed, in Wertheim theory

$$M_W = \frac{1}{2} + \frac{1}{2} \sqrt{1 + 8 \frac{\Phi}{v_d} \Delta} \quad (20)$$

where  $\Delta = V_b(e^{\beta\Delta E_S} - 1)$  and  $V_b$  is the bonding volume.<sup>75</sup> In the limit  $e^{\beta\Delta E_S} \gg 1$ , which is always valid in the  $T$ -region where chaining takes place

$$M_W = \frac{1}{2} + \frac{1}{2} \sqrt{1 + 8 \frac{V_b}{v_d} \phi e^{\beta\Delta E_S}} \quad (21)$$

The equivalence between the two expressions provides an exact definition of  $\omega$  as

$$\omega = 2 \frac{V_b}{v_d} \quad (22)$$

Although eq 19 has been derived ignoring  $O(1/M^2)$  terms in the free energy, the average chain length  $M$  can be always calculated, and this is what we do in this work, numerically locating the zero of  $\partial(\beta F/V)/\partial M = 0$ .

**B. Nematic Phase.** In the nematic phase the function  $f(\mathbf{u})$  depends explicitly on the angle between a given particle direction and the nematic axis, i.e., on the axis  $\mathbf{u}$ . The orientational distribution function  $f(\mathbf{u})$  generally depends on a set of parameters that have to be obtain through the minimization of the free energy. Also in the nematic phase we assume an exponential distribution for  $\nu(l)$ . In addition, we consider an angular distribution function  $f(\mathbf{u})$  with the form proposed by Onsager,<sup>54</sup> i.e.:

$$f(\mathbf{u}) = f_O(\mathbf{u}) = \frac{\alpha}{4\pi \sinh \alpha} \cosh(\alpha \cos \theta) \quad (23)$$

where  $\theta$  is the angle between the particle and the nematic axis and the system is supposed to have azimuthal symmetry around such axis.

In view of the analytical expression for the excluded volume  $\nu_{excl}$  for cylinders, we assume the following form for the  $\nu_{excl}$  of two DNADs averaged over the solid angle using the one parameter ( $\alpha$ ) dependent orientational distribution function  $f_O(\mathbf{u})$  defined in eq 23:

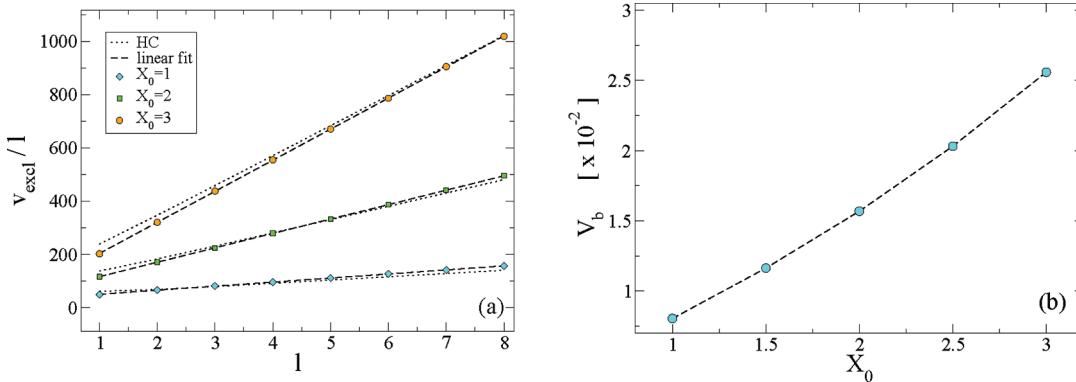
$$\nu_{excl}(l, l', \alpha) = 2 \left[ A_N(\alpha) + v_d k_N(\alpha) \frac{l + l'}{2} + B_N(\alpha) X_0^2 l' \right] \quad (24)$$

If we insert eqs 24 and 16 into eq 3, we obtain after some algebra:

$$\begin{aligned} \frac{\beta F}{V} = & \hat{\sigma}_0 - \rho(\beta\Delta E_S + \sigma_b)(1 - M^{-1}) \\ & + \eta(\phi) \left[ B_N(\alpha) X_0^2 + \frac{k_N(\alpha)}{M} v_d \right. \\ & \left. + \frac{A_N(\alpha)}{M^2} \right] \rho^2 + \frac{\rho}{M} \left( \ln \left[ \frac{v_d \rho}{M} \right] - 1 \right) \\ & - \rho \ln M + \rho \ln(M-1) \frac{M-1}{M} \end{aligned} \quad (25)$$

where  $\hat{\sigma}_0 \equiv \sum_l \sigma_0(l) \nu(l)$ .

**C. Phase Coexistence.** Using the free energy functionals in eqs 18 and 25 the phase boundaries, i.e.  $\phi_N = v_d \rho_N$  and  $\phi_I = v_d \rho_I$  of the isotropic–nematic transition can be straightforwardly calculated by minimizing the free energy with respect to the average chain lengths in the isotropic and nematic phases, i.e.  $M_I$  and  $M_N$ , and  $\alpha$ . We also require that the isotropic and nematic phases have the same pressure, i.e.,  $P_I = P_N$  and the same chemical potential  $\mu_I = \mu_N$ . These conditions require numerically solving the following set of equations:



**Figure 2.** (a) Excluded volume of two chains of length  $l$  calculated numerically as a function of  $l$  for  $X_0 = 1, 2, 3$ . Dashed lines are fits to eq 27. Dotted lines are excluded volumes for chains of HCs calculated using eq 7. (b) Bonding volume as a function of aspect ratio  $X_0$ .

$$\begin{aligned} \frac{\partial}{\partial M_I} F_{\text{iso}}(\rho_I, M_I) &= 0 \\ \frac{\partial}{\partial M_N} F_{\text{nem}}(\rho_N, M_N, \alpha) &= 0 \\ \frac{\partial}{\partial \alpha} F_{\text{nem}}(\rho_I, M_I, \alpha) &= 0 \\ P_I(\rho_I, M_I) &= P_N(\rho_N, M_N, \alpha) \\ \mu_I(\rho_I, M_I) &= \mu_N(\rho_N, M_N, \alpha) \end{aligned} \quad (26)$$

#### IV. CALCULATION OF FREE ENERGY PARAMETERS

The theory illustrated in the previous section requires the calculation of several parameters,  $V_b$ ,  $k_b$ ,  $A_b$ ,  $B_b$ ,  $k_N$ ,  $A_N$ ,  $B_N$ , and  $l_p$ . Since an explicit calculation of these parameters is very unlikely for superquadrics in the following we describe simple methods to calculate them numerically. For example, the calculation of the excluded volume between aggregates and the calculation of the bonding volume require the evaluation of complicated integrals, which can be estimated with a Monte Carlo method.<sup>77–79</sup> The general idea behind Monte Carlo is that such complicated integrals can be calculated by generating a suitable distribution of points in the domain of integration.

**A. Excluded Volume in the Isotropic Phase.** In the isotropic phase,  $v_{\text{excl}}(l, l')$  can be written as reported in eq 15. If  $l = l'$

$$v_{\text{excl}}(l, l) = 2A_I + 2k_I v_d l + 2B_I X_0^2 l^2 \quad (27)$$

Hence, from a numerical evaluation of  $v_{\text{excl}}(l, l)$  for several  $l$  values (whose detailed procedure is described in Appendix X) it is possible to estimate  $A_b$ ,  $k_b$ , and  $B_b$ . Figure 2a shows  $v_{\text{excl}}(l, l)/l$  vs  $l$ . A straight line properly describes the data for all  $X_0$  values, suggesting that  $A_b \approx 0$ . From a linear fit one obtains  $2B_I X_0^2$  (slope) and  $2k_I v_d$  (intercept). We also show in Figure 2a, the excluded volume of chains of HCs calculated using the Onsager expression reported in eq 7. Although Onsager formula is a good approximation, eq 27 provides a better representation of the “exact” (i.e., numerically estimated) excluded volume for SQs.

**B. Calculation of the Bonding Volume.** The bonding volume  $V_b$  can be calculated numerically by performing a Monte Carlo calculation of

$$V_b = \int \theta(-\Delta E_S - u_{SW} - V_{HC}) \, d\mathbf{r} \, d\Omega_1 \, \Omega_2 \quad (28)$$

where  $V_{HC} = V_{HC}(\mathbf{r}, \Omega_1, \Omega_2)$  is the hard core part of the interaction potential and  $\theta(x)$  is the Heaviside step function; i.e.,  $\theta(x) = 1$  if  $x \geq 0$  or 0 otherwise. The details of the numerical integration are reported in Appendix X. The resulting values of  $V_b$  for different  $X_0$  are shown in Figure 2b.  $V_b$  grows with  $X_0$ , an effect introduced by the different rounding of the SQ surface close to the bases. Indeed, as shown in Figure 1, on increasing  $X_0$  the base surface is more rounded and such different rounding offers a different angular width over which bonds can form. This effect will also reflect in the  $X_0$  dependence of the persistence length of the self-assembled chains, as it will be discussed in details in subsection IVE. For HCs, the aspect ratio  $X_0$  does not impact the bonding angle and so the bonding volume would be constant.

The values of  $\sigma_b(V_b)$  (the loss in entropy of forming a bond) calculated using eq 22 are around  $-6$  and thus they are comparable with the studied values of  $\beta \Delta E_S$  ( $\Delta E_S = 5.56, 6.67$ , and  $8.33$ ).

**C. Excluded volume in the nematic phase.** The excluded volume  $v_{\text{excl}}(l, l, \alpha)$  between two aggregates of equal length  $l$  can be calculated using the procedure illustrated previously for the isotropic case with the only difference that now monomers are inserted with an orientation extracted from the Onsager angular distribution defined in eq 23.

To numerically estimate  $A_N(\alpha)$ ,  $k_N(\alpha)$ , and  $B_N(\alpha)$  we specialize eq 24 to the case of  $l = l' = 2$ ,  $l = l' = 3$ ,  $l = l' = 4$ , and numerically evaluate  $v_{\text{excl}}(2, 2, \alpha)$ ,  $v_{\text{excl}}(3, 3, \alpha)$  and  $v_{\text{excl}}(4, 4, \alpha)$  for several values of  $\alpha$ . Inverting eq 24 allows us to express  $A_N(\alpha)$ ,  $k_N(\alpha)$  and  $B_N(\alpha)$  as a function of  $v_{\text{excl}}(2, 2, \alpha)$ ,  $v_{\text{excl}}(3, 3, \alpha)$  and  $v_{\text{excl}}(4, 4, \alpha)$  as explained in detail in Appendix C.

**D. Estimate of the Orientational Entropy in the Nematic Phase.** We propose to model the orientational entropy in the nematic phase using the following approximate expression proposed by Odijk<sup>45</sup> (other possibilities can be found in refs 80 and 81)

$$\begin{aligned} \hat{\sigma}_o^{od} = \sum_{l=1}^{l=\infty} v(l) \left\{ \ln \alpha + \frac{(\alpha - 1)l}{6l_p} \right. \\ \left. + \frac{5}{12} \ln \left[ \cosh \left( \frac{(\alpha - 1)l}{5l_p} \right) \right] - \frac{19}{12} \ln 2 \right\} \end{aligned} \quad (29)$$

Unfortunately, eq 29 is hardly tractable in the minimization procedure required to evaluate the equilibrium free energy and hence the two following expressions are often preferred

$$\begin{aligned}\sigma_o^{RC}(l) &= \ln(\alpha) - 1 + \frac{\alpha - 1}{6l_p} l \quad \alpha l \ll l_p \\ \sigma_o^{FC}(l) &= \ln(\alpha/4) + \frac{\alpha - 1}{4l_p} l \quad \alpha l \gg l_p\end{aligned}\quad (30)$$

which can be obtained by inserting the Onsager orientational function  $f_O(\mathbf{u})$  in eq 10 and which are highly accurate in the limit of "rigid chains" (RC) and "flexible chains" (FC).

While in the case of fixed length polymers, the knowledge of the persistence length selects one of the two expressions, in the case of equilibrium polymers, different chain lengths will contribute differently to the orientational entropy. In particular, when the chain length distribution is rather wide, it is difficult to assess if the RC (chosen in ref 46) or the FC (chosen in ref 47) limits should be used. To overcome the numerical problem, still retaining both the RC and the FC behaviors, we use the following expression for  $\hat{\sigma}_o$ :

$$\begin{aligned}\hat{\sigma}_o = & \sum_{l=1}^{l_0-1} \nu(l) \left\{ [\ln(\alpha) - 1] + \frac{\alpha - 1}{6l_p} l \right\} + \\ & \sum_{l=l_0}^{\infty} \nu(l) \left\{ \ln(\alpha/4) + \frac{\alpha - 1}{4l_p} l \right\}\end{aligned}\quad (31)$$

in which the contribution of chains of size  $l_0$  is treated with the RC expression while the contribution of longer chains enters with the FC expression. We pick  $l_0$  by requiring the maximum likelihood between eq 31 and eq 29 in the relevant  $M\alpha$  domain. The procedure is quite straightforward: given a appropriate domain in the  $(M,\alpha)$  plane (e.g.,  $5 < M < 150$  and  $5 < \alpha < 50$ ), we chose a grid  $G$  within this domain and we determined the value of  $l_0$  which minimizes the following function  $T(l_0)$ :

$$T(l_0) \equiv \min_G |\hat{\sigma}_o^{od} - \hat{\sigma}_o| \quad (32)$$

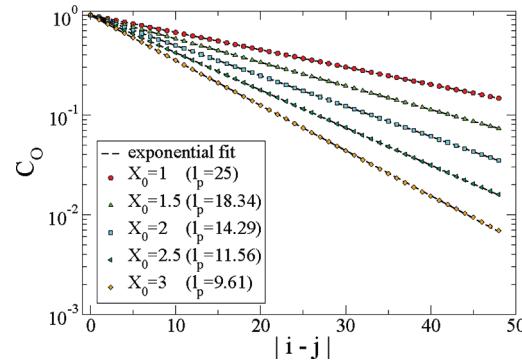
Although  $l_0$  weakly depends on  $l_p$  (i.e.,  $l_0 = l_0(l_p)$ ), we find that the value  $l_0 \approx 9$  is appropriate for most studied cases.

**E. Estimate of Persistence Length.** In order to estimate the persistence length, entering in eq 31, we randomly build chains according to the procedure described in Appendix B. We estimate the "chain persistence length"  $l_p$  by evaluating following spatial correlation function:

$$C_O(|i-j|) \equiv \sum_{i,j} \langle \hat{\mathbf{x}}(i) \cdot \hat{\mathbf{x}}(j) \rangle \quad (33)$$

where  $i,j$  label two monomers along the chain ( $i = 0$  is the first monomer at chain end) and  $\hat{\mathbf{x}}(i)$  is a unit vector directed along  $x$ -axis of the monomer (i.e., their axis of symmetry, see Figure 1), that coincides with the direction along which the two attractive sites lie.  $\langle \dots \rangle$  denotes an average over the whole set of independent chains which has been generated.

In Figure 3 we plot  $C_O(|i-j|)$  for all aspect ratios studied. All correlations decay following an exponential law, whose characteristic scale is identified as the persistence length (in unit of monomer). In the  $X_0$  range explored,  $10 < l_p < 25$ . The more elongated monomers have a smaller persistence length.



**Figure 3.** Spatial correlation function  $C_O(|i-j|)$  (see text for its definition) calculated by generating random chains of 50 monomers for aspect ratios  $X_0 = 1, 1.5, 2, 2.5, 3$ . Dashed lines are fits to the functional form  $C_O(|i-j|) = \exp[-|i-j|/l_p]$ . From these fits, the chain persistence length  $l_p$  can be estimated (see legend).

The  $X_0$  dependence of  $l_p$  arises from the different roundness of the bases (implicit in the use of SQ), as discussed in the context of the bonding volume and in Figure 1.

## V. RESULTS AND DISCUSSION

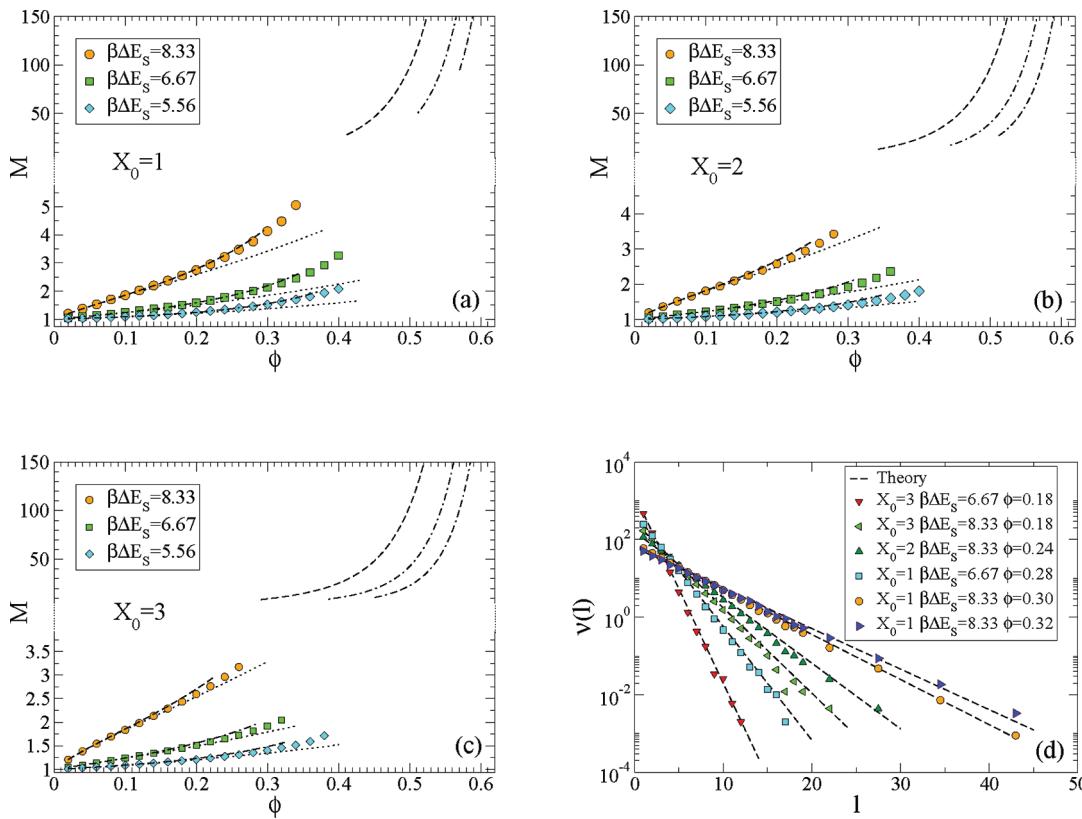
In this section we compare results from simulations with theoretical calculations based on the theory discussed in section III

**A. Isotropic Phase.** Parts a–c of Figure 4 show the packing fraction dependence of  $M$  for  $X_0 = 1, 2, 3$  for all temperatures investigated. The dashed curves are calculated by minimizing the isotropic free energy in eq 18 with respect to  $M$  using the values of  $V_b$ ,  $k_b$ , and  $B_I$  obtained in subsection IVA without any fitting. Up to volume fractions around  $\phi \approx 0.20$  the agreement between theoretical and numerical results is quite good for all cases considered. Above this volume fraction the theoretical predictions start deviating appreciably, a discrepancy that we attribute at moderate and high  $\phi$  to the inaccuracy of the Parsons decoupling approximation. We also plot in Figure 4, parts a–c, as dotted lines the predictions based on an Onsager-like theory, i.e. setting the Parsons-Lee factor  $\eta(\phi)$  equal to 1 in eq 3. At low volume fractions the approximation  $\eta(\phi) = 1$  does not affect the quality of the results but above  $\phi \approx 0.20$  the use of Parsons decoupling approximation seems to better capture the behavior of  $M(\phi)$ . In Figure 4d, we report the aggregate size distribution  $\nu(l)$  as obtained from both simulation and theory, the latter calculated according to eq 18 with  $M$  obtained by minimization of the isotropic free energy. As expected, the aggregate size in the isotropic phase is exponential. These results suggest that a reasonable first principles description of the isotropic phase is provided by the free energy of eq 18, when the parameters of the model are properly evaluated.

**B. Nematic Phase.** On increasing  $\phi$  the system transforms into a LC phase. We estimate the degree of nematic ordering by evaluating the largest eigenvalue  $S$  of the order tensor  $\mathbf{Q}$ , whose components are:

$$Q_{\alpha\beta} = \frac{1}{N} \sum_i \frac{3}{2} \langle (\mathbf{u}_i)_\alpha (\mathbf{u}_i)_\beta \rangle - \frac{1}{2} \delta_{\alpha\beta} \quad (34)$$

where  $\alpha\beta \in \{x,y,z\}$ , and the unit vector  $(\mathbf{u}_i(t))_\alpha$  is the component  $\alpha$  of the orientation (i.e., the symmetry axis) of particle  $i$  at time  $t$ . A nonzero value of  $S$  signals the presence of orientational order in the system and it can be found not only



**Figure 4.** (a–c) Average chain length  $M$  against  $\phi$  for  $X_0 = 1, 2, 3$  for the three studied  $\beta\Delta E_S$  values. Symbols are MC results in the isotropic phase. Dashed lines are theoretical predictions, obtained by minimizing the free energy in eqs 18 (isotropic phase, bottom part of the figures) and 25 (nematic phase, upper part of the figures). Different types of dashed lines refer to different stacking energies. The dashed lines are limited by the region of stability of the isotropic and nematic phases. Dotted lines are theoretical predictions according to Onsager second virial approximation, i.e. setting  $\eta(\phi) = 1$ . Please note the change of scale along  $y$ -axis. (d) MC results for the aggregate size distributions (colored symbols) for several state points in the isotropic phase and corresponding theoretical predictions (dashed lines).

in the nematic phase but also in partially ordered phases as columnar and smectic phases. Since in this article we focus only on the nematic phase, to verify that the simulated state points are not partially ordered we calculate, following ref 46, the three-dimensional pair distribution function  $g(\mathbf{r})$  defined as

$$g(\mathbf{r}) = \frac{1}{\rho N} \left\langle \sum_{i=1}^N \sum_{j \neq i} \delta(\mathbf{r} - (\mathbf{r}_i - \mathbf{r}_j)) \right\rangle \quad (35)$$

where  $\delta(\mathbf{x})$  is the Dirac delta function. We calculate the  $g(\mathbf{r})$  in a reference system with the  $z$ -axis parallel to the nematic director. Figure 5 shows  $g(x,y,0)$  and  $g(0,y,z)$ , which correspond, respectively, to the correlations in a plane perpendicular to the nematic director and in a plane containing it for a given nematic state point ( $X_0 = 2, \phi = 0.38, \beta\Delta E_S = 8.33$ ). The  $g(x,y,0)$  is found to be isotropic, ruling out the possibility of a columnar or crystal phase (no hexagonal symmetry is indeed present). The  $g(0,y,z)$  reflects the orientational ordering along the nematic direction and rules out the possibility of a smectic phase (no aligned sequence of peaks are present<sup>46</sup>). Figure 5c also shows a snapshot of the simulated system at the same state point.

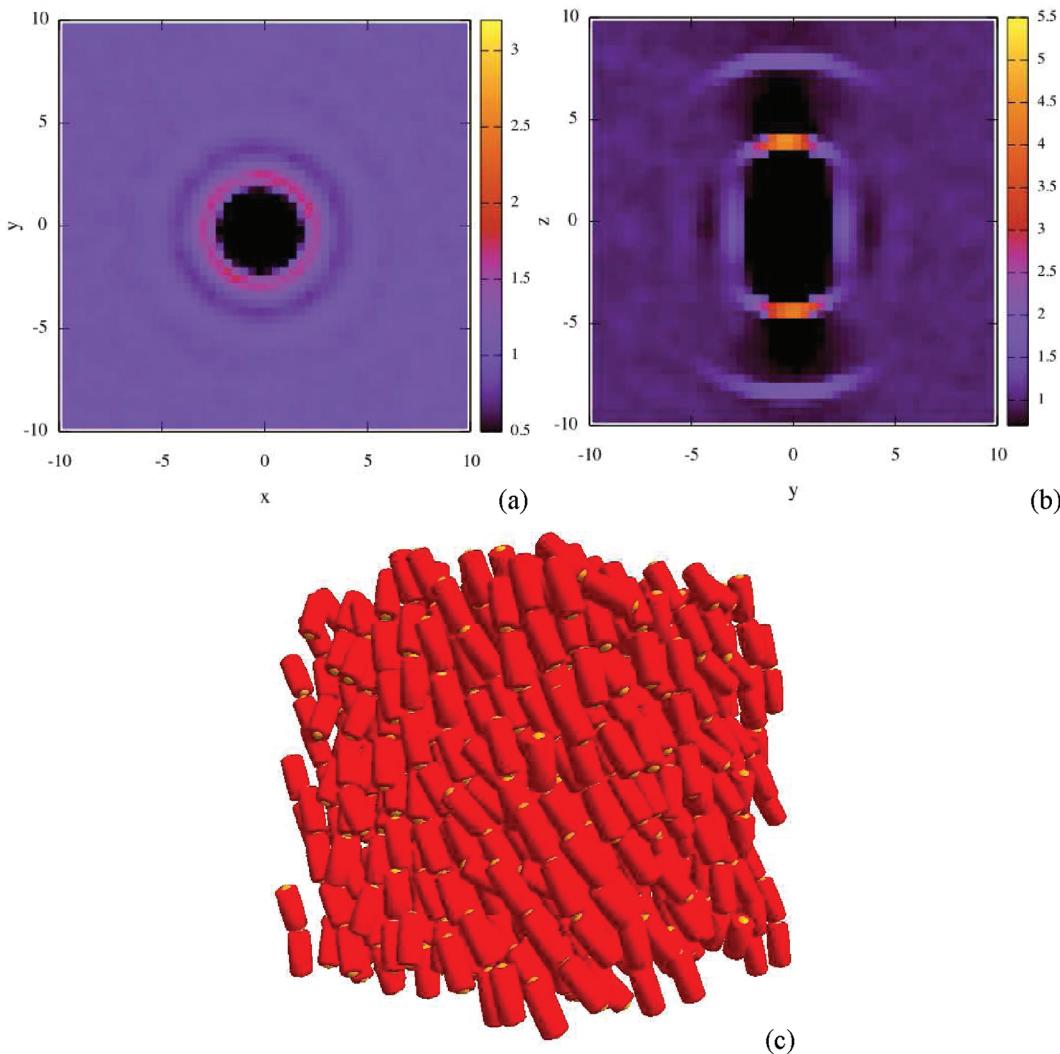
In what follows, we have systematically calculated and inspected  $g(\mathbf{r})$  to verify that all state points having a value of  $S$  large enough to be considered nematic are indeed translationally isotropic, i.e. with no translational order.

Figure 6 shows the nematic order parameter and the average chain length  $M$  calculated from simulations as well as with the

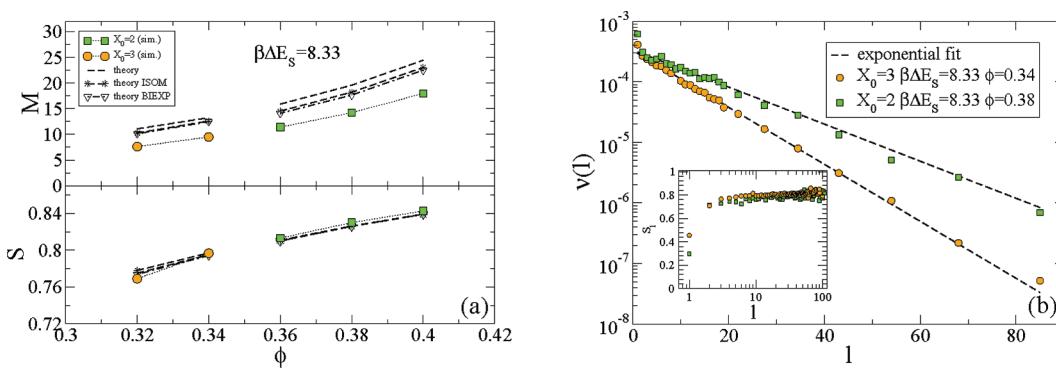
theoretical methodology described previously for two different aspect ratios at  $\beta\Delta E_S = 8.33$ . The theoretical value for  $S$  is obtained according to:

$$S(\alpha) = \int 2\pi \frac{3 \cos^2 \theta - 1}{2} f_O(\theta; \alpha) \sin \theta d\theta \quad (36)$$

Figure 6a shows that the nematic order parameter is very well captured by the theory, while the average chain length shows a clear disagreement between theory and simulations, again suggesting that the error introduced by the Parsons decoupling approximation, previously discussed in the case of the isotropic phase at large packing fractions, is enhanced by the further increase in  $\phi$ . Another possible source of error could arise from the hypothesis that the aggregate size distribution is also exponential in the nematic phase. To test this hypothesis Figure 6b shows the aggregate size distributions at two different state points. In all cases, the distributions are not a single exponential. This phenomenon has been already observed and discussed by Lu and Kindt,<sup>47</sup> who described the distribution with two exponential decays of  $\nu(l)$  with the exponential decay of short chains extending up to  $l \approx 50$ . They took into account such a biexponential nature of the distribution to better reproduce the isotropic–nematic phase boundaries<sup>82</sup> in their theoretical approach. In the present case, only very short chains (not to say only the monomers), fall out of the single exponential decay. To test if the different decay reflects a different orientational ordering of the small aggregates compared to long chains, we



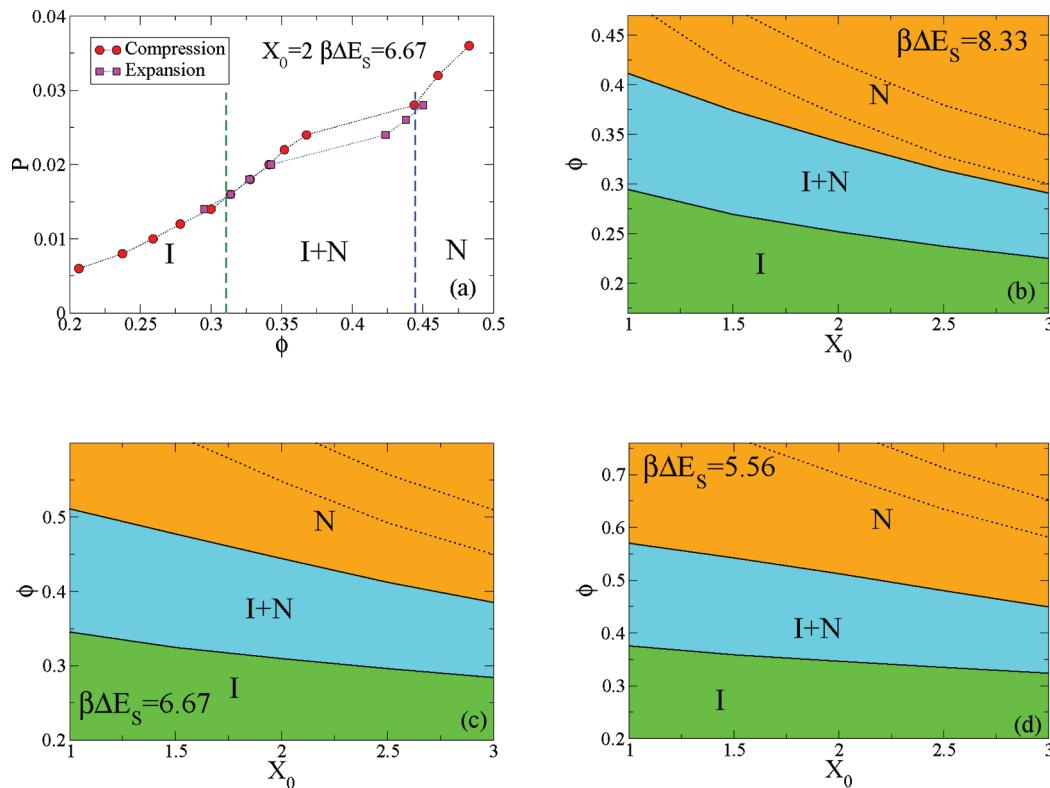
**Figure 5.** Plot of  $g(x,y,0)$  (a) and  $g(0,y,z)$  (b) where the  $z$ -axis is chosen parallel to the nematic director for  $X_0 = 2$ ,  $\phi = 0.38$ , and  $\beta\Delta E_S = 8.33$ . (c) Example of nematic configurations at the same state point.



**Figure 6.** (a) Average chain length and nematic order parameter  $S$  for several nematic state points. Dashed lines with stars (theory ISOM) and down triangles (theory BIEXP) are improved theoretical predictions (see text for details). (b) Aggregate size distribution for two state points ( $X_0 = 2, \phi = 0.38, \beta\Delta E_S = 8.33$ ) and ( $X_0 = 3, \phi = 0.34, \beta\Delta E_S = 8.33$ ). Circles are numerical results and dashed lines are exponential fits. The inset shows the chain length dependent nematic order parameter  $S_l$  for the same state points.

follow ref 82 and evaluate the length-dependent nematic order parameter  $S_l$ , that is the nematic order parameter calculated for each population of aggregates of size  $l$ . The results, reported in the inset of Figure 6b, show that  $S_l$  is around 0.7–0.8 for all aggregates sizes except for  $l = 1$ , i.e., except for monomers.

To assess how much the theoretical predictions are affected by the assumption of a single exponential decay (and of the associated identity of  $S$  for all chains), we evaluate the correction of the free energy functional in eq 25 for the two following cases:



**Figure 7.** (a) Equation of state ( $P$  vs  $\phi$ ) calculated compressing an isotropic initial configuration (squares) or expanding an initial nematic configurations (circles). Vertical dashed lines show the theoretical predictions for the phase boundaries. (b–d) Coexistence regions predicted from theoretical calculations for the three stacking energies values  $\beta\Delta E_S$  investigated. Dotted lines are phase boundaries calculated within Onsager (second virial) approximation. For each  $X_0$ , the appropriate persistence length (see Figure 3), has been selected.

- (i) We retain a single exponential distribution  $\nu(l)$  but monomers are assumed to be isotropic, while all other chains are nematic.
- (ii) Monomers are assumed to be isotropic but we also assume a bimodal chain length distribution  $\nu(l)$ , i.e.

$$\nu(l) = \begin{cases} \rho_I & l = 1 \\ \nu_N(l, \rho - \rho_I, M_N) & l > 1 \end{cases} \quad (37)$$

where  $\rho_I$  is the number density of monomers, which is calculated by free energy minimization, and

$$\begin{aligned} \nu_N(l, \rho - \rho_I, M_N) \\ = \frac{(M_N - 1)^{l-2} M_N^{1-l}}{M_N + 1} (\rho - \rho_I) \end{aligned} \quad (38)$$

Note also that the average chain length for the above choice of  $\nu(l)$  is

$$M = \frac{\rho}{\rho_I + \frac{\rho - \rho_I}{M_N + 1}} \quad (39)$$

In case i, we exclude from the calculation of the orientational entropy the monomers and take into account the fact that monomers are isotropic in the excluded volume calculation. The revised free energy can be thus written as

$$\begin{aligned} \frac{\beta F}{V} = & \hat{\sigma}_0^* - \rho(\beta\Delta E_S + \sigma_b)(1 - M^{-1}) \\ & + \eta(\phi) \left[ B_N(\alpha)X_0^2 + \frac{k_N(\alpha)}{M}\nu_d \right. \\ & \left. + \frac{A_N(\alpha)}{M^2} \right] \rho^2 - \beta\Delta f_N + \frac{\rho}{M} \\ & \left( \ln \left[ \frac{\nu_d \rho}{M} \right] - 1 \right) - \rho \ln M + \rho \\ & \ln(M - 1) \frac{M - 1}{M} \end{aligned} \quad (40)$$

where

$$\begin{aligned} \sigma_0^* = & \sum_{l=2}^{l=l_0-1} \nu(l) \left\{ [\ln(\alpha) - 1] + \frac{\alpha - 1}{6l_p} l \right\} + \\ & \sum_{l=l_0}^{l=\infty} \nu(l) \left\{ \ln(\alpha/4) + \frac{\alpha - 1}{4l_p} l \right\} \end{aligned} \quad (41)$$

and

$$\begin{aligned} \beta\Delta f_N = & \eta(\phi) \left\{ \nu_d [k_N(\alpha) - k_I] \left( \frac{1}{M^2} + \frac{1}{M^3} \right) \right. \\ & \left. + 2X_0^2 [B_N(\alpha) - B_I] \frac{1}{M^2} + 2 \frac{A_N(\alpha)}{M^3} \right\} \rho^2 \end{aligned} \quad (42)$$

In case ii, after some more cumbersome calculations, one obtains the following expression for the free energy:

$$\begin{aligned} \frac{\beta F}{V} = & \hat{\sigma}_0^* - (\rho - \rho_I)(\beta \Delta E_S + \sigma_b) \\ & \left(1 - \frac{1}{M_N + 1}\right) + \eta(\phi) \left[ B_N(\alpha) X_0^2 \right. \\ & + \frac{k_N(\alpha)}{M_N + 1} v_d + \frac{A_N(\alpha)}{(M_N + 1)^2} \left. \right] (\rho - \rho_I)^2 \\ & + \eta(\phi) \left\{ k_I v_d \left[ \rho_I^2 + \rho_I(\rho - \rho_I) \right. \right. \\ & \left. \left. \left(1 + \frac{1}{M_N + 1}\right) \right] + X_0^2 B_I \right. \\ & \left. [\rho_I^2 + 2\rho_I(\rho - \rho_I)] \right\} + \frac{M_N - 1}{M_N + 1} (\rho - \rho_I) \\ & \ln(M_N - 1) - M_N \frac{\rho - \rho_I}{M_N + 1} \ln(M_N) \\ & - \frac{\rho - \rho_I}{M_N + 1} \ln(M_N + 1) \\ & + \rho_I [\ln(v_d \rho_I) - 1] + \frac{\rho - \rho_I}{M_N + 1} \\ & \left. \{ \ln[v_d(\rho - \rho_I)] - 1 \} \right\} \end{aligned} \quad (43)$$

In passing, we note that eq 43 reduces to eq 40 if we assume a single exponential distribution, i.e., if we set  $\rho_I = \rho/M^2$  and  $M_N = M$ , and if we neglect terms  $O(1/M^4)$ .

Minimizing the free energy expressions in eqs 43 and 40 results in an improved estimate for the average chain length and nematic order parameter (Figure 6, see curves labeled by (i) ISOM and (ii) BIEXP). The new estimates slightly improve over the previous ones, suggesting once more that the leading source of error in the present approach, as well in all previous ones, has to be found in the difficulty of properly handling the higher order terms in the virial expansion.

**C. Phase Coexistence.** NPT-MC simulations only provide a rough estimate of the location of phase boundaries, since these simulations can be affected by the hysteresis associated with the metastability of the coexisting phases. It is thus only possible to bracket the region of coexistence, by selecting the first isotropic state point on expansion runs which started from a nematic configuration and the first nematic state point on compression runs started from an isotropic configuration. We performed NPT-MC simulations for  $X_0 = 2$  and  $\beta \Delta E_S = 6.67$  over a wide range of pressures  $P$  for a system of 1000 SQs. The resulting equation of state is shown in Figure 7a. As expected, a clear hysteresis is observed, which allows us to detect only some overestimated boundaries for the isotropic–nematic transition. The same figure also reports the theoretical estimates of the transition. The theoretical critical pressure is smaller than the numerical one, resulting in a more extended region of coexistence than observed numerically. Comparing the values of the pressure predicted by the theory with the simulation values, we notice that the main error arises from the pressure of the nematic phase, which is underestimated. Parts b–d of

Figure 7 show the predicted phase diagram for several values of  $\beta \Delta E_S$  as a function of the aspect ratio. On increasing  $\beta \Delta E_S$  (i.e., decreasing  $T$  or increasing the stacking energy), there is a small decrease of  $\phi_I$  and a significant decrease of  $\phi_N$ , resulting in an overall decrease of the I–N coexistence region. Such trends can be understood in terms of the increase of the average chain length resulting from the increase of  $\beta \Delta E_S$ . The theoretical values for the average chain length at the nematic–isotropic coexistence are shown in Figure 4, parts a–c. Along the  $\phi_I$  transition line,  $M$  ranges from 2 to 4. On the contrary, along the  $\phi_N$  transition line, the  $M$  values are larger and depend on aspect ratio and stacking energy.

As expected, both  $\phi_I$  and  $\phi_N$  decrease on increasing  $X_0$ . In parts b–d of Figure 7, we also plot the phase boundaries calculated within the Onsager (second virial) approximation, i.e., by setting  $\eta(\phi) = 1$  in our free energy functional. It is clear from these plots that neglecting higher order terms in the virial expansion results in a significant overestimate of the isotropic–nematic transition volume fraction. Hence the Parsons–Lee approximation captures, albeit approximatively, the contribution of higher order virial coefficients on increasing the volume fraction, which are neglected in the Onsager approximation.

Finally, we recall that in our model the persistence length  $l_p$  depends on the aspect ratio as discussed in subsection IV.E.

## VI. COMPARISON WITH EXPERIMENTS

References 29 and 31 report the critical concentrations ( $c$ ), in mg/mL, for the I–N transition of blunt-ended DNAD. These experimental data can be transformed into volume fractions once the relevant properties of DNAD are known (DNAD molecular weight  $m_D = 660N_b$  Da, diameter  $D \approx 2$  nm, and length  $L = N_b/3$  nm, where  $N_b$  is the number of bases in the sequence). The number density  $\rho$  of DNADs is related to the mass concentration

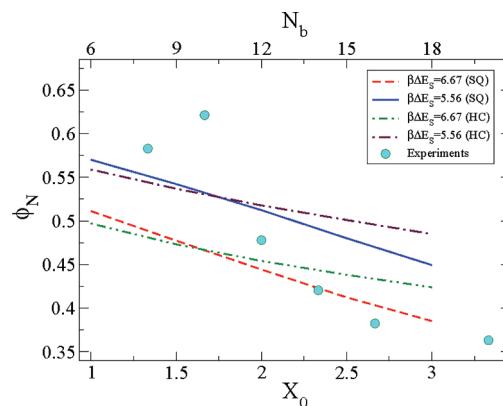
$$\rho = \frac{c}{m_D} \quad (44)$$

Since  $v_d = LD^2\pi/4$  is the volume of a DNAD, the volume fraction can be expressed as:

$$\varphi = \rho v_d = \frac{cLD^2\pi}{4m_D} \quad (45)$$

Data in refs 29 and 31 suggest that blunt-end duplexes of equal length but different sequences may have different transition concentrations. As discussed in refs 31 and 83, this phenomenon can be attributed to the slight differences in B-DNA helical conformation resulting from the difference in sequences. These differences induce some curvature in the DNAD aggregates, in turn enhancing the transition concentration. Indeed, sequences that are known to form straight double helices order into the N phase at lower concentrations. Therefore, for each oligomer length in the range 8–16 bases, we selected the lowest transition concentration among the ones experimentally determined, since these would be closest to the symmetric monomers considered in the model. Such values have been reported in Figure 8 as a function of base number  $N_b$  (top axis) and as a function of  $X_0$  (bottom axis). Apart for  $N_b = 12$ , for which a large number of sequences have been studied, the transition concentrations for the other  $N_b$  values would probably be corrected to lower values if a larger number of sequences were experimentally explored. We would expect this to be particularly true for the shortest sequences, in which the effect of bent helices could be more relevant.

In the experiments, DNADs are in a water solution with counterions resulting from the dissociation of the ionic groups



**Figure 8.** Critical volume fractions  $\Phi_N$  as a function of aspect ratio  $X_0$  (or equivalently  $N_b$ ) from theoretical calculations for SQs and HCs (for  $\beta\Delta E_S = 6.67$  and  $\beta\Delta E_S = 5.56$ ) and for experiments<sup>29</sup> (circles).

of the phosphate-sugar chain. Given the high DNA concentration necessary for the formation of the N phase, corresponding to concentration of nucleobases in the 1 M range, the ionic strength simply provided by the natural counterions is large enough to effectively screen electrostatic interactions between DNADs. This becomes less true for the longest studied sequences, for which the transition concentration is lower. We hence decided to perform a small number of test experiments on the  $N_b = 20$  oligomers with a double purpose: (i) determine more accurately the transition concentration value for this compound and (ii) test the effect of varying the ionic strength predicted by the model here described. With respect to a fully screened DNAD, where electrostatic repulsion can be neglected, a partly screened DNAD has a larger effective volume, thus filling a larger volume fraction of the solution, and a smaller axial ratio  $X_0$ , since electrostatic repulsion is equal in all directions. Therefore, adding salt would bring about two competing effects: the reduction in particle volume, which enhances the concentration needed to reach the I–N phase boundary, and the growth of  $X_0$ , which could favor the nematic ordering even at lower concentrations.

The behavior expected upon increasing the salt concentration can be obtained combining two elements: (i) From eq 45, we can deduce the following relation between the critical concentration  $c_N$  and the critical volume fraction  $\phi_N$ :

$$c_N = \phi_N(X_0) \frac{4m_D}{LD^2\pi} \quad (46)$$

(ii) The phase diagrams of Figure 7b–d indicates that  $\phi_N(X_0)$  depends weakly on  $X_0 = L/D$ , i.e.,  $\phi_N(X_0) = \phi_N^0$ , where  $\phi_N^0$  is constant.

Hence the theory introduced in the present paper predicts that a reduction of DNAD effective volume due to the addition of salt (i.e., a decrease of  $LD^2$  in eq 46 leads to an overall increase of the concentration required for N ordering.

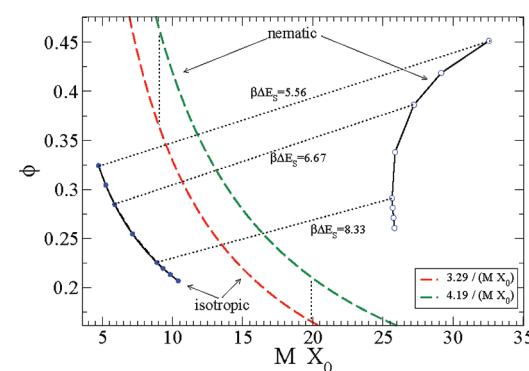
We have measured the transition concentration of the self-complementary 20mer CGCGAAATTTCGCG, a sequence whose I–N transition at room temperature was previously measured and determined to be  $c_N \approx 200$  mg/mL.<sup>29</sup> With the same method, based on the measurement of the refractive index of the solution, we determined the I–N transition concentration at room temperature at three different ionic strengths. The values we obtained are  $c_N \approx 215$  mg/mL (no added salt),  $c_N \approx 320$  mg/mL (0.8 M NaCl), and  $c_N \approx 380$  mg/mL (1.2 M NaCl). The data indicate that the onset of the nematic ordering in solutions of 20mers is indeed sensitive to the ionic strength, and that the

transition concentration grows upon increasing the amount of salt, as expected on the basis of our theoretical calculations for the present model. In Figure 8, we display the transition volume fraction derived by the transition concentration measured for 1.2 M NaCl. At this ionic strength, the total concentration of  $\text{Na}^+$  ions (from the assumed fully dissociated oligomers + added with the salt) is about the same as the one resulting from counterions dissociated oligomers in the more concentrated solutions of shorter (8–12 mers) oligomers.

Figure 8 compares the experimentally determined transition volume fractions with the values calculated from the SQ model for  $\beta\Delta E_S = 6.67$  and  $\beta\Delta E_S = 5.56$ . Although the experimental data are noisy, they fall in the range  $\Delta E_S \approx 5–7$  (in units of  $k_B T$ ). Despite all the simplifying assumptions and despite the experimental uncertainty, the results in Figure 8 provide a reasonable description of the  $X_0$  dependence of  $\phi_N$ . In this figure we also reported the theoretical volume fractions calculated for HCs with a fixed bonding value equal to  $V_b(X_0 = 2) \approx 0.0157$ . Chains of HCs have a persistence length which does not depend on  $X_0$ . It is clear that the estimated  $\Delta E_S$  for HCs is unchanged with respect to SQ, i.e., the phase diagram of our model is not significantly affected by the  $X_0$  dependence of the persistence length (at least in the investigated windows).

In comparing the model with the experimental results, it is necessary to take note of the fact that the stacking energy between nucleobases, and thus the interaction energy  $\Delta E_S$  between DNAD, is temperature dependent, i.e., its entropic component is relevant.<sup>35</sup> This is a general property of solvation energies and thus it is in line with the notion that stacking forces are mainly of a hydrophobic nature. Therefore, the range of values for  $\Delta E_S$  determined in Figure 8 should be compared to the values of  $\Delta G$  for the stacking interactions at the temperature at which the experiments were performed. Overall, the estimate of  $\Delta E_S$  here obtained appears as in reasonable agreement with the free energies involved in the thermodynamic stability of the DNA double helices and confirms the rough estimate that was given before (see the supporting online material associated with ref 29).

Experimental data are often compared to the original Onsager theory for monodisperse thin hard rods, approximating the polydispersity created by the aggregation process as an average aspect ratio  $MX_0$ .<sup>16,29</sup> As a guide to the interpretation of such data, we compare in Figure 9 the theoretically estimated



**Figure 9.** Isotropic–nematic coexistence lines in the average aspect ratio  $MX_0$  and packing  $\phi$  plane. Solid lines with symbols indicate theoretical predictions for  $X_0 = 3$ . Symbols along the isotropic and nematic phases at coexistence are joined by dotted lines, to indicate the change in packing and average chain length at the transition. Dashed lines indicate the Onsager original predictions, as re-evaluated in ref 37 for  $\phi_I$  and  $\phi_N$ . In this case, the tie lines (dotted) are vertical.

isotropic–nematic coexistence lines in the  $MX_0 - \phi$  plane, parametrized by the stacking energy. The figure shows also the Onsager original predictions, as re-evaluated in ref 37, in which no change of aspect ratio takes place at coexistence. Including polydispersity as a result of reversible assembling (and higher order excluded volume contributions) significantly alters the coexisting values of  $\phi$ .

## VII. CONCLUSIONS

In this study we have developed a free energy functional to calculate the isotropic–nematic phase diagram of bifunctional quasi-cylindrical monomers, aggregating into equilibrium chains. The model has been inspired by experiments on the aggregation of short DNA, which exhibits nematic phases at sufficiently high concentrations. The comparison between the theoretical predictions and the experimental results allows us to provide an estimate of the stacking energy, consistent with previous propositions.

Our approach is quite general, parameter free and not restricted to particular shapes. Once the shape of the monomer and geometry of the sticky sites are defined, the model dependent quantities, like bonding volume, persistence length and excluded volume can be calculated either theoretically, as for HCs, or numerically, as done in the present case resorting to a MC integration technique. Using these quantities the phase diagram can be evaluated without any fit or adjustment of the parameters and we do not need to perform any simulation to calculate these quantities.

We provide techniques to evaluate the bonding volume and the excluded volume, which enters into our formalism via the Parsons–Lee decoupling approximation. We build on previous work, retaining the discrete aggregate size description of ref 47 and the Parsons–Lee factor for the excluded volume contribution proposed in ref 46. With respect to previous approaches, we (i) explicitly account for the entropic and energetic contributions associated with bond formation, and (ii) we do not retain any adjustable fit parameter.

The resulting description of the isotropic phase is rather satisfactory and quantitative up to  $\phi \approx 0.2$ . The description of the nematic phase partially suffers from some of the approximations made in deriving the free energy functional. More specifically, several signatures point toward the failure of the Parsons decoupling approximation in the  $\phi$  range typical of the nematic phase. While there is a sufficient understanding of the quality of such approximation for monodisperse objects,<sup>63,64,66,68,69,73</sup> work needs to be done to assess the origin of the failure of this approximation in the equilibrium polymer case and to propose improvements.

We finally note that the model introduced here does not consider the azimuthal rotations of each monomer around its axis. This neglect is adequate when the aggregation does not entail constraints in the azimuthal freedom of the monomers. This is the case of base stacking, in which the angular dependence of the stacking energy is arguably rather small. However, this is not the case of DNA interacting through the pairing of overhangs and of the LC ordering of RNA duplexes. Because of its A-DNA-type structure, the terminal paired bases of RNA duplexes are significantly tilted with respect to the duplex axis, thus establishing a link between the azimuthal angle of the aggregating duplexes and the straightness of the aggregate even in the case of blunt ended duplexes. However, with minor modifications the model here introduced could become suitable to include these additional situations. The

limiting factor in developing such extensions is the lack of knowledge to quantify the azimuthal constraints implied by these interactions. This situation, as well as the effects of off-axis components of the end-to-end interduplex interactions, will be explored in a future work.

## APPENDIX A

Here we provide a justification for the use of Parsons decoupling approximation in the case of linear chains poly disperse in length (with distribution  $\nu(l)$ ), based on the extension of Onsager's second-virial theory to mixtures of nonspherical hard bodies proposed in ref 84. The contribution  $F_{excl}$  to the free energy due to excluded volume interactions between chains can be written if we neglect intrachain interactions:<sup>84,85</sup>

$$\frac{\beta F_{excl}}{V} = \frac{\rho}{6} \int_0^\rho d\rho' \int d\mathbf{r} \int d\Omega_1 d\Omega_2 \sum_{ll'} \frac{\nu(l)\nu(l')}{\rho^2} g_{ll'}(\mathbf{r}, \Omega_1, \Omega_2) f(\Omega_1) f(\Omega_2) \mathbf{r} \cdot \nabla_{\mathbf{r}} V_{HC}(\mathbf{r}, \Omega_1, \Omega_2) \quad (47)$$

where  $\mathbf{r}$  is the distance between the centers of mass of the two chains 1 and 2,  $\Omega_1 = \{\mathbf{u}_1^1, \dots, \mathbf{u}_l^1\}$  and  $\Omega_2 = \{\mathbf{u}_1^2, \dots, \mathbf{u}_{l'}^2\}$  are the orientations of the two chains, where  $\mathbf{u}_i^\alpha$  is the orientation of monomer  $i$  belonging to chain  $\alpha = 1, 2$ ,  $g_{ll'}(\mathbf{r}, \Omega_1, \Omega_2)$  is the molecular radial distribution function of the mixture, which represents the correlations between two chains of length  $l$  and  $l'$ , whose relative distance is  $\mathbf{r}$  and which have orientations  $\Omega_1$  and  $\Omega_2$  respectively,  $V_{HC}(\mathbf{r}, \Omega_1, \Omega_2)$  is the hard-core part of the interaction potential and  $f(\Omega_\alpha)$  is the angular distribution function of chain  $\alpha$ . We note that in eq 47 the integration in  $\rho'$  is performed keeping fixed all the parameters related to  $f(\Omega_\alpha)$ . Neglecting intrachain interactions is equivalent to ignore self-overlaps of chains, an assumption which is appropriate if chain length is not much greater than its persistence length and the chains can be considered nonextensible.

Parsons decoupling approximations in this case accounts to putting:

$$g_{ll'}(\mathbf{r}, \Omega_1, \Omega_2) = g_{ll'}^{HS}[r/\sigma_{ll'}(\hat{r}, \Omega_1, \Omega_2)] \quad (48)$$

where  $g_{ll'}^{HS}$  is the radial distribution function of a mixture of hard spheres and  $\sigma_{ll'}(\hat{r}, \Omega_1, \Omega_2)$  is an angle-dependent range parameter which depends on chain lengths  $l$  and  $l'$ . If the pair interaction is of the special form

$$V_{HC}(\mathbf{r}, \Omega_1, \Omega_2) = V_{HC}[r/\sigma_{ll'}(\hat{r}, \Omega_1, \Omega_2)] \quad (49)$$

noting that  $\mathbf{r} \cdot \nabla_{\mathbf{r}} = r(\partial)/(\partial r)$ , eq 47 becomes

$$\frac{\beta F_{excl}}{V} = \frac{\rho}{6} \int_0^\rho d\rho' \int d\hat{r} d\Omega_1 d\Omega_2 \sum_{ll'} \frac{\nu(l)\nu(l')}{\rho^2} \int dr r^3 g_{ll'}^{HS}(r/\sigma_{ll'}) f(\Omega_1) f(\Omega_2) \frac{\partial V_{HC}(r/\sigma_{ll'})}{\partial r} \quad (50)$$

With the substitution  $y = r/\sigma_{ll'}$  from eq 50 one obtains:

$$\begin{aligned} \frac{\beta F_{excl}}{V} &= \frac{\rho}{2} \sum_{ll'} \frac{1}{3} \frac{\nu(l)\nu(l')}{\rho^2} \int_0^\rho d\rho' \int d\hat{r} d\Omega_1 \\ &\quad d\Omega_2 \int dy y^3 \frac{\partial V_{HC}(y)}{\partial y} g_{ll'}^{HS}(y) f(\Omega_1) f(\Omega_2) \sigma_{ll'}^3 \end{aligned} \quad (51)$$

The derivative of  $V_{HC}$  is a delta function hence we need only to evaluate the value of  $g_{ll'}^{HS}(y)$  at contact (i.e.,  $y = 1^+$ ) and eq 51 becomes:

$$\begin{aligned} \frac{\beta F_{excl}}{V} &= \frac{\rho}{2} \sum_{ll'} \frac{\nu(l)\nu(l')}{\rho^2} \int d\rho' g_{ll'}^{HS}(1^+) \int d\hat{r} d\Omega_1 \\ &\quad d\Omega_2 f(\Omega_1) f(\Omega_2) \frac{\sigma_{ll'}^3}{3} \end{aligned}$$

This expression tends to Parsons–Lee's expression when the system is monodisperse ( $\nu(l) = \rho\delta_{l,l}$ ). In the specific case of spherical particles  $\sigma_{ll'}(\hat{r}, \Omega_1, \Omega_2) = \sigma(\hat{r}, \Omega_1, \Omega_2)$  and

$$\begin{aligned} \sum_{ll'} \nu(l)\nu(l') f(\Omega_1) f(\Omega_2) \int d\hat{r} d\Omega_1 d\Omega_2 \\ \frac{1}{3} \sigma(\hat{r}, \Omega_1, \Omega_2) \\ = \frac{4\pi}{3} \sigma^3 \end{aligned} \quad (52)$$

i.e., the excluded volume of two spheres of diameter  $\sigma$ . Hence we are allowed to make the identification:

$$\nu_{excl}(l, l') = \int d\hat{r} d\Omega_1 d\Omega_2 \frac{1}{3} f(\Omega_1) f(\Omega_2) \sigma_{ll'}^3 \quad (\hat{r}, \Omega_1, \Omega_2) \quad (53)$$

and write:

$$\frac{\beta F_{excl}}{V} = \frac{\rho}{2} \sum_{ll'} \frac{\nu(l)\nu(l')}{\rho^2} \left[ \int d\rho' g_{ll'}^{HS}(1^+) \right] \nu_{excl}(l, l') \quad (54)$$

We note that the identification made in eq 53 can be also further justified using the same reasonings given in section III. As discussed in ref 84, a possible expression for  $g_{ll'}^{HS}$  is the one derived by Boublik,<sup>86</sup> which generalizes the Carnahan–Starling relation<sup>87</sup> for pure hard spheres to the case of mixtures, i.e.

$$\begin{aligned} g_{ll'}^{HS}(1^+) &= \frac{1}{1 - \zeta_3} + \frac{3\zeta_2}{(1 - \zeta_3)^2} \frac{\hat{\sigma}_{ll}\hat{\sigma}_{l'l'}}{\hat{\sigma}_{ll} + \hat{\sigma}_{l'l'}} \\ &+ \frac{2\zeta_2^2}{(1 - \zeta_3)^3} \frac{(\hat{\sigma}_{ll}\hat{\sigma}_{l'l'})^2}{(\hat{\sigma}_{ll} + \hat{\sigma}_{l'l'})^2} \end{aligned} \quad (55)$$

where  $\hat{\sigma}_{ll}$  is the diameter of an hard sphere corresponding to a chain of length  $l$  and  $\zeta_n = (\pi/6) \sum \nu(l) \hat{\sigma}_{ll}^n$ . To map the system of polydisperse chains onto the equivalent mixture of hard spheres we need an expression for  $\hat{\sigma}_{ll}$ . According to ref 84, the simplest choice is to consider spheres having the same volume of the corresponding linear chain of length  $l$ , i.e.

$$\nu_d = \frac{\pi}{6l} \hat{\sigma}_{ll}^3 \quad (56)$$

where we recall that  $\nu_d$  is the volume of a monomer. Although in principle we could use eq 54 together with eqs 55 and 56 to calculate the free energy contribution due to the excluded volume between particles, if we make the further assumption that

$$g_{ll'}^{HS}(1^+) \approx g^{HS}(1^+) \quad (57)$$

i.e., if we approximate the radial distribution function of the hard spheres mixture at contact with that of a monodisperse system of hard spheres having the same total volume fraction (i.e., setting in eq 55  $\hat{\sigma}_{ll'} = \hat{\sigma}$  with  $M\nu_d = (\pi/6)\sigma^3$ ), we finally obtain

$$\frac{\beta F_{excl}}{V} = \frac{\eta(\varphi)}{2} \sum_{ll'} \nu(l)\nu(l') \nu_{excl}(l, l') \quad (58)$$

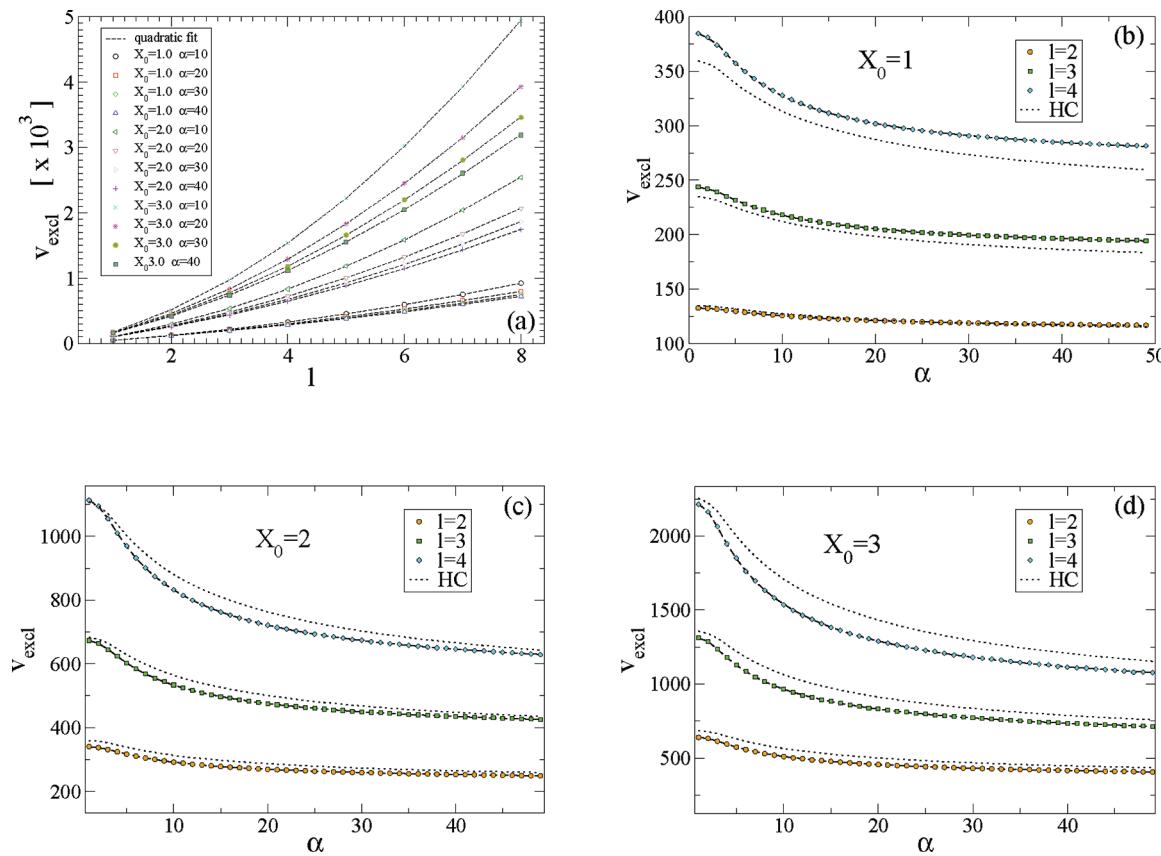
where we used the Carnahan–Starling expression for  $g^{HS}(1^+; \rho')$  and we performed the integration in  $\rho'$ . Equation 58 is exactly the expression for the contribution to the free energy due to steric repulsion which we used in section III. In summary according to the above derivation we argue that eq 58 can be not accurate at high volume fractions due to the approximations made in eqs 48 (i.e., the Parsons decoupling approximation) and 57. Within the present treatment, eq 58 is also not appropriate for chains with  $l \gg l_p$  because, as already noted, chain self-overlaps can be significant and the hard body pair potential  $V_{HC}$  does not have the special form assumed in eq 49.

We finally note that the approximation made in eq 57 can be avoided if one resorts to eq 54 instead of eq 58, although the required free energy calculations would become much more complicated. Anyway we verified for the isotropic phase that employing eq 54 instead of eq 58 does not provide any appreciable improvement in the present case.

## APPENDIX B

The procedure, which we adopted to calculate the excluded volume  $\nu_{excl}$ , recalls the one described in ref 77, except that we generate new configurations without any bias technique and by assuming an angular distribution for monomer orientations, which is uniform in the isotropic phase and the Onsager orientational distribution function in the nematic phase. Our procedure in the isotropic phase consists in fact in performing  $N_{att}$  attempts of inserting two chains of length  $l$  in a box of volume  $V$  as described in the following list:

- 1 Set the counter  $N_{ov} = 0$
- 2 Build first chain of length  $l$  randomly, according to the following procedure:
  - (a) Insert a first randomly oriented monomer.
  - (b) Insert a monomer  $M$  bonded to a free site  $S$  on chain ends ( $S$  can be chosen randomly among the two free sites of the partial chain). The orientation of  $M$  will be random and its position will be chosen randomly within the available bonding volume between  $M$  and  $S$ . The bonding volume between  $M$  and  $S$  is defined as the volume corresponding to all possible center of mass positions of  $M$  with  $M$  bonded to  $S$ .
- (c) If the number of monomer inserted is  $l$  terminate otherwise go to point 1
- (d) where the first monomer inserted is placed in the center of the box and it is oriented with its attractive sites parallel to the  $x$ -axis.
- 3 Build a second chain of length  $l$ , where the first monomer inserted is placed randomly within the simulation box with a random orientation.



**Figure 10.** (a) Excluded volume of two chains of length  $l$  as a function of chain length for the nematic cases  $\alpha = 10, 20, 30, 40$  and three different aspect ratios  $X_0 = 1, 2, 3$ . (b–d) Excluded volume in the nematic phase calculated numerically as a function of  $\alpha$  for two chains of equal length  $l$ , where  $l = 2, 3, 4$ , composed of monomers with  $X_0 = 1, 2, 3$ . Dotted lines represent the excluded volume calculated with Onsager formula reported in eq 7.

- 4 Increase  $N_{ov}$  by 1 if two monomers belonging to different chains overlap and the two chains are either not self-overlapping or forming a closed loop.
- 5 If the number of attempts is less than  $N_{att}$  go to point 2; otherwise, terminate.

Then  $v_{\text{excl}}$  can be calculated as follows:

$$v_{\text{excl}} = \frac{N_{ov}}{N_{att}} V \quad (59)$$

A reasonable choice for the total number of attempts is  $N_{att} = 10^6$ . In a similar fashion one can also calculate the bonding volume<sup>75</sup> between two monomers. In this case one monomer is kept fixed in the center of the simulation box and the other one is inserted with random position and orientation for a total of  $N_{att}$  attempts. The bonding volume will be:

$$V_b = \frac{N_{bond}}{4N_{att}} V \quad (60)$$

where the factor 4 accounts for the fact that two particles can form 4 different possible bonds and  $N_{bond}$  is the number of times that the two monomers were bonded after a random insertion. Finally with the same procedure used to calculate the excluded volume in the isotropic phase we can evaluate the excluded volume in the nematic phase. The only difference is that now monomers have to be inserted with an orientation extracted from the Onsager angular distribution defined in eq 23, so that the excluded volume depends also on the parameter

$\alpha$ . Again if  $N_{ov}$  is the number of times that two monomers belonging to different aggregates overlap and  $N_{att}$  is the total number of attempts then we have:

$$v_{\text{excl}}(l, l, \alpha) = \frac{N_{ov}}{N_{att}} V \quad (61)$$

## APPENDIX C

In this appendix, we explain how to calculate the parameters  $A_N(\alpha)$ ,  $k_N(\alpha)$ , and  $B_N(\alpha)$  of the nematic free energy functional. As a preliminary step we check that  $v_{\text{excl}}(l, l, \alpha)$  for a fixed value of  $\alpha$  is a second order polynomial of  $l$  and  $l'$  as assumed in eq 5. In Figure 10a, we plot  $v_{\text{excl}}(l, l, \alpha)$  as a function of  $l$  for different values of  $\alpha$  and  $X_0$ ,  $v_{\text{excl}}(l, l, \alpha)$  can be well represented by a parabolic function, in agreement with eq 5.

We start by observing that the  $\alpha$  dependence of  $A_N(\alpha)$ ,  $k_N(\alpha)$ , and  $B_N(\alpha)$  in the case of hard cylinders following the Onsager distribution can be expanded in powers of  $\alpha^{-1/2}$  as

$$\begin{aligned} A_N(\alpha) &= c_{00} + \frac{c_{01}}{\alpha^{1/2}} + \frac{c_{02}}{\alpha} + \frac{c_{03}}{\alpha^{3/2}} + \frac{c_{04}}{\alpha^2} \\ k_N(\alpha) &= c_{10} + \frac{c_{11}}{\alpha^{1/2}} + \frac{c_{12}}{\alpha} + \frac{c_{13}}{\alpha^{3/2}} + \frac{c_{14}}{\alpha^2} \\ B_N(\alpha) &= c_{20} + \frac{c_{21}}{\alpha^{1/2}} + \frac{c_{22}}{\alpha} + \frac{c_{23}}{\alpha^{3/2}} + \frac{c_{24}}{\alpha^2} \end{aligned} \quad (62)$$

where  $c_{ij}$  are the elements of the  $3 \times 4$  matrix  $\mathbf{C}$ . In the case of cylinders, some of the  $c_{ij}$  vanishes.<sup>45</sup> We assume here that the same  $\alpha$  dependence holds for SQs.

In view of this result the covolume as a function of  $l$  and  $\alpha$  can be expressed as

$$\begin{aligned} v_{\text{excl}}^{(\text{fit})}(\alpha; X_0, l) = & d_{l0} + \frac{d_{l1}}{\alpha^{1/2}} + \frac{d_{l2}}{\alpha} + \frac{d_{l3}}{\alpha^{3/2}} \\ & + \frac{d_{l4}}{\alpha^2} \end{aligned} \quad (63)$$

where  $d_{lp}$ , for  $p = 0 \dots 4$  are fitting parameters. Parts b-d of Figure 10 show the numerical calculation of the covolume varying  $\alpha$  for three particular aspect ratios ( $X_0 = 1, 2, 3$ ), together with fits to the functional form of eq 63 and the excluded volume calculated using the Onsager expression for HC (see eq 7). Onsager formula as in the isotropic case provides only an approximate description of SQs "exact" (i.e., numerically estimated) excluded volume.

The good quality of the fits (reduced  $\chi^2$  is always much less than 1 for all fits) suggests that retaining terms up to  $O(1/\alpha^2)$  is to the present level of accuracy of our calculations absolutely appropriate.

From these fits we can estimate the matrix  $\mathbf{C}$  needed to evaluate the free energy in the nematic phase for each  $X_0$ . If we define in fact the following matrix  $\mathbf{P}$  and the vectors  $\mathbf{q}_p$ , with  $p = 0 \dots 4$  as follows:

$$\mathbf{P} = \begin{pmatrix} 1 & l_a & l_a^2 \\ 1 & l_b & l_b^2 \\ 1 & l_c & l_c^2 \end{pmatrix} \mathbf{q}_p = \begin{pmatrix} d_{l_ap} \\ d_{l_bp} \\ d_{l_cp} \end{pmatrix} \quad (64)$$

where  $l_a, l_b$ , and  $l_c$  are three different chain lengths for which we calculated the  $v_{\text{excl}}$  as a function of  $\alpha$ , then we can calculate the matrix elements of  $\mathbf{C}$  in the following way:

$$2 \begin{pmatrix} c_{0p} \\ v_d c_{1p} \\ X_0^2 c_{2p} \end{pmatrix} = \mathbf{P}^{-1} \mathbf{q}_p \quad (65)$$

## ACKNOWLEDGMENTS

C.D.M. and F.S. acknowledge support from ERC (226207-PATCHYCOLLOIDS). T.B. acknowledges support from the Italian MIUR Ministry (grant PRIN-2008F3734A). The authors thank Prof. S. Kumar for his careful reading of the manuscript and G. Zanchetta for the support in the experimental tests.

## REFERENCES

- (1) Hamley, I., *Introduction to Soft Matter*; Wiley & Sons: New York, 2007.
- (2) Glotzer, S. C. *Science* **2004**, *306*, 419.
- (3) Whitesides, G. M.; Boncheva, M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 4769.
- (4) Workum, V.; Douglas, J. *Phys. Rev. E* **2006**, *73*, 031502.
- (5) Mirkin, C.; Letsinger, R.; Mucic, R.; Storhoff, J. *Nature* **1996**, *382*, 607.
- (6) Manoharan, V. N.; Elsesser, M. T.; Pine, D. J. *Science* **2003**, *301*, 483.

- (7) Cho, Y.-S.; Yi, G.-R.; Lim, J.-M.; Kim, S.-H.; Manoharan, V. N.; Pine, D. J.; Yang, S.-M. *J. Am. Chem. Soc.* **2005**, *127*, 15968.
- (8) Yi, G.; Manoharan, V. N.; Michel, E.; Elsesser, M. T.; Yang, S.; Pine, D. J. *Adv. Mater.* **2004**, *16*, 1204.
- (9) Starr, F. W.; Douglas, J. F.; Glotzer, S. C. *J. Chem. Phys.* **2003**, *119*, 1777.
- (10) Starr, F. W.; Sciortino, F. *J. Phys.: Condens. Matter* **2006**, *18*, L347.
- (11) Stupp, S. I.; Son, S.; Lin, H. C.; Li, L. S. *Science* **1993**, *259*, 59.
- (12) Doye, J. P. K.; Louis, A. A.; Lin, I.-C.; Allen, L. R.; Noya, E. G.; Wilber, A. W.; Kok, H. C.; Lyus, R. *Phys. Chem. Chem. Phys.* **2007**, *9*, 2197.
- (13) Khan, A. *Curr. Opin. Colloid Interface Sci.* **1996**, *1*, 614.
- (14) van der Schoot, P.; Cates, M. *Langmuir* **1994**, *10*, 670.
- (15) Kuntz, D. M.; Walker, L. M. *Soft Matter* **2008**, *4*, 286.
- (16) Jung, J.-M.; Mezzenga, R. *Langmuir* **2010**, *26*, 504.
- (17) Lee, C. F. *Phys. Rev. E* **2009**, *80*, 031902.
- (18) Ciferri, A. *Liq. Cryst.* **2007**, *34*, 693.
- (19) Aggeli, A.; Bell, M.; Carrick, L. M.; Fishwick, C. W. G.; Harding, R.; Mawer, P. J.; Radford, S. E.; Strong, A. E.; Boden, N. *J. Am. Chem. Soc.* **2003**, *125*, 9619.
- (20) Robinson, C. *Tetrahedron* **1961**, *13*, 219.
- (21) Livolant, F.; Levelut, A. M.; Doucet, J.; Benoit, J. P. *Nature* **1989**, *339*, 724.
- (22) Merchant, K.; Rill, R. L. *Biophys. J.* **1997**, *73*, 3154.
- (23) Tombolato, F.; Ferrarini, A. *J. Chem. Phys.* **2005**, *122*, 054908.
- (24) Tombolato, F.; Ferrarini, A.; Grelet, E. *Phys. Rev. Lett.* **2006**, *96*, 258302.
- (25) Barry, E.; Beller, D.; Dogic, Z. *Soft Matter* **2009**, *5*, 2563.
- (26) Grelet, E.; Fraden, S. *Phys. Rev. Lett.* **2003**, *90*, 198302.
- (27) Tomar, S.; Green, M. M.; Day, L. A. *J. Am. Chem. Soc.* **2007**, *129*, 3367.
- (28) Minsky, A.; Shimon, E.; Frenkel-Krispin, D. *Nat. Rev. Mol. Cell. Biol.* **2002**, *3*, 50.
- (29) Nakata, M.; Zanchetta, G.; Chapman, B. D.; Jones, C. D.; Cross, J. O.; Pindak, R.; Bellini, T.; Clark, N. A. *Science* **2007**, *318*, 1276.
- (30) Zanchetta, G.; Nakata, M.; Buscaglia, M.; Clark, N. A.; Bellini, T. *J. Phys.: Condens. Matter* **2008**, *20*, 494214.
- (31) Zanchetta, G.; Giavazzi, F.; Nakata, M.; Buscaglia, M.; Cerbino, R.; Clark, N. A.; Bellini, T. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 17497.
- (32) Zanchetta, G.; Bellini, T.; Nakata, M.; Clark, N. A. *J. Am. Chem. Soc.* **2008**, *130*, 12864.
- (33) Lydon, J. J. *Mater. Chem.* **2010**, *20*, 10071.
- (34) Guckian, K. M.; Schweitzer, B. A.; Ren, R. X.-F.; Sheils, C. J.; Tahmassebi, D. C.; Kool, E. T. *J. Am. Chem. Soc.* **2000**, *122*, 2213.
- (35) Bellini, T.; Cerbino, R.; Zanchetta, G. *Top. Curr. Chem.* **2011**, *1*–55.
- (36) Budin, I.; Szostak, J. W. *Annu. Rev. Biophys.* **2010**, *39*, 245.
- (37) Vroege, G. J.; Lekkerkerker, H. N. W. *Rep. Prog. Phys.* **1992**, *55*, 1241.
- (38) Dijkstra, M.; Frenkel, D. *Phys. Rev. E* **1995**, *51*, 5891.
- (39) Khokhlov, A.; Semenov, A. *Physica* **1981**, *108A*, 546.
- (40) Khokhlov, A.; Semenov, A. *Physica* **1982**, *112A*, 605.
- (41) Wessels, P. P. F.; Mulder, B. M. *J. Phys.: Condens. Matter* **2006**, *18*, 9335.
- (42) Dennison, M.; Dijkstra, M.; van Roij, R. *Phys. Rev. Lett.* **2011**, *106*, 208302.
- (43) Wang, Z.; Kuckling, D.; Johannsmann, D. *Soft Mater.* **2003**, *1*, 353.
- (44) Chen, Z. Y. *Macromolecules* **1993**, *26*, 3419.
- (45) Odijk, T. *Macromolecules* **1986**, *19*, 2313.
- (46) Kouriabova, T.; Betterton, M.; Glaser, M. *J. Mat. Chem.* **2010**, *20*, 10366.
- (47) Lü, X.; Kindt, J. *J. Chem. Phys.* **2004**, *120*, 10328.
- (48) van der Schoot, P.; Cates, M. *Europhys. Lett.* **1994**, *25*, 515.
- (49) Chen, B.; Siepmann, J. *J. Phys. Chem. B* **2000**, *104*, 8725.
- (50) Chen, B.; Siepmann, J. *J. Phys. Chem. B* **2001**, *105*, 11275.
- (51) Wertheim, M. *J. Stat. Phys.* **1984**, *35*, 19.

- (52) Wertheim, M. *J. Stat. Phys.* **1984**, *35*, 35.  
(53) Wertheim, M. *J. Stat. Phys.* **1986**, *42*, 459.  
(54) Onsager, L.; Ann., N. Y. *Acad. Sci.* **1949**, 627.  
(55) De Michele, C. *J. Comput. Phys.* **2010**, *229*, 3276.  
(56) De Michele, C.; Gabrielli, S.; Tartaglia, P.; Sciortino, F. *J. Phys. Chem. B* **2006**, *110*, 8064.  
(57) De Michele, C.; Tartaglia, P.; Sciortino, F. *J. Chem. Phys.* **2006**, *125*, 204710.  
(58) Corezzi, S.; De Michele, C.; Zaccarelli, E.; Fioretto, D.; Sciortino, F. *Soft Matter* **2008**, *4*, 1173.  
(59) Corezzi, S.; De Michele, C.; Zaccarelli, E.; Tartaglia, P.; Sciortino, F. *J. Phys. Chem. B* **2009**, *113*, 1233.  
(60) Lee, C.; Cammon, J. M.; Rossky, P. *J. Chem. Phys.* **1984**, *80*, 4448.  
(61) Parsons, J. *Phys. Rev. A* **1979**, *19*, 1225.  
(62) Lee, S. *J. Chem. Phys.* **1987**, *87*, 4972.  
(63) Camp, P. J.; Mason, C. P.; Allen, M. P.; Khare, A. A.; Kofke, D. *A. J. Chem. Phys.* **1996**, *105*, 2837.  
(64) Wensink, H.; Lekkerkerker, H. *Mol. Phys.* **2009**, *107*, 2111.  
(65) Varga, S.; Szalai, I. *Mol. Phys.* **2000**, *98*, 693.  
(66) Wensink, H.; Vroege, G.; Lekkerkerker, H. *J. Phys. Chem. B* **2001**, *105*, 10610.  
(67) Gámez, F.; Merkling, P.; Lago, S. *Chem. Phys. Lett.* **2010**, *494*, 45.  
(68) Cinacchi, G.; Mederos, L.; Velasco, E. *J. Chem. Phys.* **2004**, *121*, 3854.  
(69) McGrother, S. C.; Williamson, D. C.; Jackson, G. *J. Chem. Phys.* **1996**, *104*, 6755.  
(70) Wensink, H. H.; Vroege, G. J.; Lekkerkerker, H. N. W. *J. Chem. Phys.* **2001**, *115*, 7319.  
(71) Szabolcs, V.; Galindo, A.; Jackson, G. *Mol. Phys.* **2003**, *101*, 817.  
(72) Galindo, A.; Haslam, A. J.; Varga, S.; Jackson, G.; Vanakaras, A. G.; Photinos, D. J.; Dunmur, D. A. *J. Chem. Phys.* **2003**, *119*, 5216.  
(73) Cuetos, A.; Martnez-Haya, B.; Lago, S.; Rull, L. *Phys. Rev. E* **2007**, *75*, 061701.  
(74) Williamson, D.; Jackson, G. *J. Chem. Phys.* **1998**, *108*, 10294.  
(75) Sciortino, F.; Bianchi, E.; Douglas, J. F.; Tartaglia, P. *J. Chem. Phys.* **2007**, *126*, 194903.  
(76) Jackson, G.; Chapman, W. G.; Gubbins, K. E. *Mol. Phys.* **1988**, *65*, 1.  
(77) Fynnewever, H.; Yethiraj, A. *J. Chem. Phys.* **1998**, *108*, 1636.  
(78) Movahed, H. B.; Hidalgo, R. C.; Sullivan, D. E. *Phys. Rev. E* **2006**, *73*, 032701.  
(79) Ibarra-Avalos, N.; Gil-Villegas, A.; Richa, A. M. *Mol. Simul.* **2007**, *33*, 505.  
(80) DuPré, D. B.; Yang, S.-J. *J. Chem. Phys.* **1991**, *94*, 7466.  
(81) Hentschke, R. *Macromolecules* **1990**, *23*, 1192.  
(82) Lü, X.; Kindt, J. *J. Chem. Phys.* **2006**, *125*, 054909.  
(83) Frezza, E.; Tombolato, F.; Ferrarini, A. *Soft Matter* **2011**, *7*, 9291.  
(84) Malijevský, A.; Jackson, G.; Varga, S. *J. Chem. Phys.* **2008**, *129*, 144504.  
(85) Honnell, K. G.; Hall, C. K.; Dickman, R. *J. Chem. Phys.* **1987**, *87*, 664.  
(86) Boublík, T. *J. Chem. Phys.* **1970**, *53*, 471.  
(87) Carnahan, N. F.; Starling, K. E. *J. Chem. Phys.* **1969**, *51*, 635.