Eur. Phys. J. E (2019) **42**: 71
DOI 10.1140/epje/i2019-11836-5

THE EUROPEAN PHYSICAL JOURNAL E

Regular Article

Free energy of conformational isomers: The case of gapped DNA duplexes

Alberto Giacomo Orellana[1] and Cristiano De Michele[2,a]

[1] SISSA, Scuola Internazionale Superiore Studi Avanzati, Via Bonomea 265, I-34136 Trieste, Italy
[2] Dipartimento di Fisica, "Sapienza" Università di Roma, P.le A. Moro 2, 00185 Roma, Italy

Received 11 April 2019 and Received in final form 19 April 2019
Published online: 7 June 2019

Abstract. Liquid-crystalline phases in all-DNA systems have been extensively studied in the past and although nematic, cholesteric and columnar mesophases have been observed, the smectic phase remained elusive. Recently, it has been found evidence of a smectic-A ordering in an all-DNA system, where the constituent particles, which are gapped DNA duplexes, resemble chain-sticks. It has been argued that in the smectic-A phase these DNA chain-sticks should be folded as a means to suppress aggregate polydispersity and excluded volume. Nevertheless, if initial crystalline configurations are prepared *in silico* with gapped DNA duplexes either fully unfolded or fully folded by carrying out computer simulations one can end up with two different phases having at the same concentration and temperature the majority of gapped DNA duplexes either folded or unfolded. This result suggests that these two phases have a small free energy difference, since no transition is observed from one to the other within the simulation time span. In the present manuscript, we assess which of these two phases is thermodynamically stable through a suitable protocol based on thermodynamic integration. Our method is rather general and it can be used to discriminate stable states from metastable ones of comparable free energy.

1 Introduction

Liquid-crystalline phases of both short [1–10] and long [11–22] B-DNA duplexes in aqueous solution have been widely investigated in the past. Mesophases, which include nematic, cholesteric and columnar hexagonal phases, emerge at sufficiently high concentration and for short duplexes they are understood in terms of end-to-end stacking interactions between duplex terminals, which induce the formation of semi-flexible linear aggregates [1]. Among liquid crystals (LC) based on DNA the smectic phase was elusive. Even if it was initially proposed by Strzelecka *et al.* [23] for 50 nm DNA duplexes, right after Livolant *et al.* [15,24] disproved this result by showing that columnar ordering preempts the formation of smectic layering in this system. More recently, it has been provided both experimental and numerical evidence of this phase by studying a slightly different system, *i.e.* short gapped B-DNA duplexes (gDNADs) [5]. The experimental evidence of a smectic-A type phase has been obtained by introducing a single-stranded DNA spacer between two double-stranded duplexes. The resulting DNA construct is thus a gDNAD which resembles a nunchaku, *i.e.* a chain-stick, which is the typical weapon of Japanese martial arts, such as Kung-fu and Ju-Jitsu. Monte Carlo simulations were carried out by using a suitable coarse-grained model, which has been developed building on a model proposed some years ago for short DNA duplexes [25–28]. Employing this coarse-grained model —which will be illustrated in sect. 2.1— in ref. [5] MC simulations in the constant pressure ensemble have been carried out. The initial configuration was generated as a crystalline lattice of fully folded gDNADs —*i.e.* employing a particular conformational isomer— in a orthorhombic lattice and the equation of state shows a clear evidence of smectic-A phase where gDNADs are folded (in ref. [5] such a phase has been called *Sm*-fA, where f stands for folded). In principle, a smectic phase, where gapped duplexes are mostly unfolded (*Sm*-uA), can also be attained, although this phase should be entropically unfavored, due to the entropic cost of building up a layered structure of linear chains with length polydispersity and the larger excluded volume of the unfolded conformation. Other conformations than the fully folded and fully unfolded ones can be excluded since they are not compatible with smectic ordering. In ref. [5] to address the thermodynamic stability of the *Sm*-fA phase with respect to the *Sm*-uA, MC simulations carried out under thermodynamic conditions appropriate for the emergence

[a] e-mail: cristiano.demichele@uniroma1.it

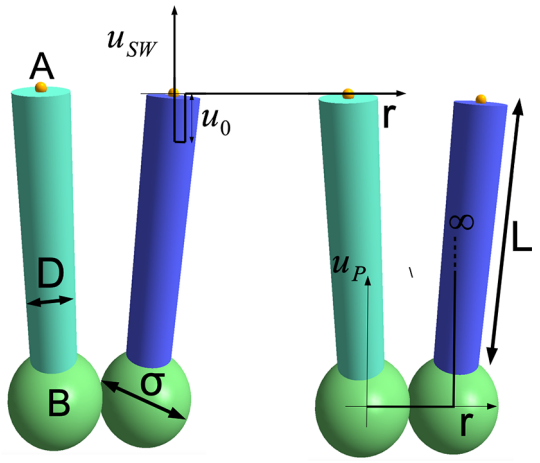

Fig. 1. Coarse-grained model of a gDNA used in the MC simulations.

of the smectic phase and starting with a random distribution of folding angles were performed. These simulations, although a fully equilibrated final state could not be achieved within the very long simulation time span, due to the very slow folding kinetics, suggests that the final equilibrium state is the *Sm*-fA. In this paper, we provide an unambiguous and conclusive proof of the thermodynamic stability of the *Sm*-fA phase resorting to thermodynamic integration [29,30] between the *Sm*-fA phase and the *Sm*-uA. For the crystalline phase thermodynamic integration can be used to calculate the absolute value of the free energy by switching on a suitable external potential, which couples the system to an Einstein crystal [31–34]. Here, building on some ideas first proposed in refs. [35,36], we follow a different approach where the system is coupled to external potentials which preserve the symmetry of the system and which allow to go from a partially ordered state (a mesophase) to a fully disordered phase without any intervening phase transition. Our method could be used both to calculate absolute values of free energy or free energy differences. In the former case one needs to evaluate the free energy of a gas to have free energy absolute values.

2 Methods

2.1 Model

As already noted, the model of a gDNAD which we used in the present manuscript is identical to the one employed in ref. [5], but we will describe it in the following for the sake of completeness. A gDNAD is modeled in a coarse-grained manner as two irreversibly bonded hard cylinders with length $L = 16\,\text{nm}$ and thickness $D = 3\,\text{nm}$ (so that the aspect ratio $X_0 = L/D = 5.33$), as shown in fig. 1. The diameter of the cylinders has been chosen larger than the steric diameter of B-DNA which is around 2 nm to account for gDNA electrostatic repulsion. The latter choice can be justified by noting that, under typical experimental conditions for observing DNA mesophases, the Debye screening length is expected to be of few angstroms, *i.e.* electrostatic interactions are strongly screened.

Each cylinder is decorated with two interacting sites labelled with A and B in fig. 1. Site B is the center of the orange sphere of diameter σ and it is located along the symmetry axes of the cylinder at a distance $L/2 + \sigma/2$ from its geometrical center. If r is the distance between two B sites belonging to the same gDNA (for B sites of distinct gDNAs $u_P = 0$), these sites interact through the following potential u_P:

$$u_P(\mathbf{r}) = \begin{cases} 0, & \text{if } r < \sigma, \\ \infty, & \text{otherwise.} \end{cases} \tag{1}$$

Site A is the center of the small green sphere of diameter δ and it is placed on the symmetry axis of the cylinder at a distance equal to $L/2 + 0.15D/2$ from the center. Sites A belonging to distinct gDNA interact through the following square well potential u_{SW}:

$$u_{SW}(\mathbf{r}) = \begin{cases} u_0, & \text{if } r < \delta, \\ 0, & \text{otherwise,} \end{cases} \tag{2}$$

where $\delta = 0.16D$. Sites A model hydrophobic attraction between DNA terminals and their geometry is the same one used in ref. [37]. The attraction strength is set to $\beta u_0 = 8$ which is the value used in ref. [5] and which provides a stacking free energy in line with values previously determined from the phase behaviour [25,26,37] and cholesteric properties [27] of self-assembling ultrashort DNA duplexes. We note that the attraction between A sites, *i.e.* between DNA terminal ends, is the same end-to-end stacking force of hydrophobic origin, which drives the formation of liquid crystal phases of short DNA duplexes [1,25,3].

Interacting sites B serve to model the single-stranded DNA which holds together the two double-stranded parts of the gDNAD. We note that the choice for the geometry and interaction potential of the B sites ensures full flexibility of the gDNAD. In the following we will make use of reduced temperature $T^* = Tk_B/u_0$ and reduced pressure $P^* = \beta P v_{hc}$, where v_{hc} is the volume of each cylinder forming the gDNAD.

2.2 Monte Carlo simulations

Monte Carlo simulations were performed in the NPT ensemble. The system is composed of $N = 840$ gDNADs which have been modeled as discussed in sect. 2.1. The simulation box of volume V was orthorhombic and the three edges are allowed to change independently in the MC simulation to ease the accommodation of smectic and crystal phases. We applied standard periodic boundary conditions. In the following $\rho = N/V$ will be the number density of gDNADs and $\phi = 2\rho v_{hc}$ will be the system volume fraction.

We built two equations of states by carrying out a series of NPT-MC simulations at different pressures starting from two difference crystalline configurations, where in

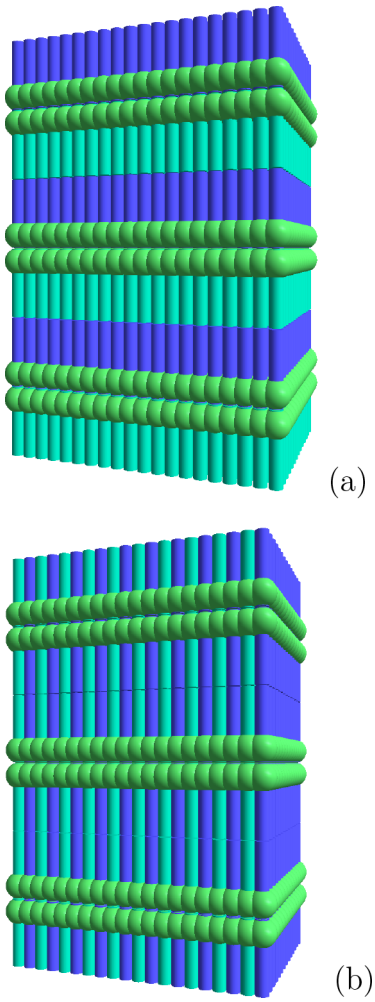

Fig. 2. Starting configurations for MC simulations: (a) crystal of unfolded gDNADs and (b) crystal of folded gDNADs. Two hard cylinders belonging to same duplex are colored in different colors (blue and green) to make folded and folded configurations more apparent.

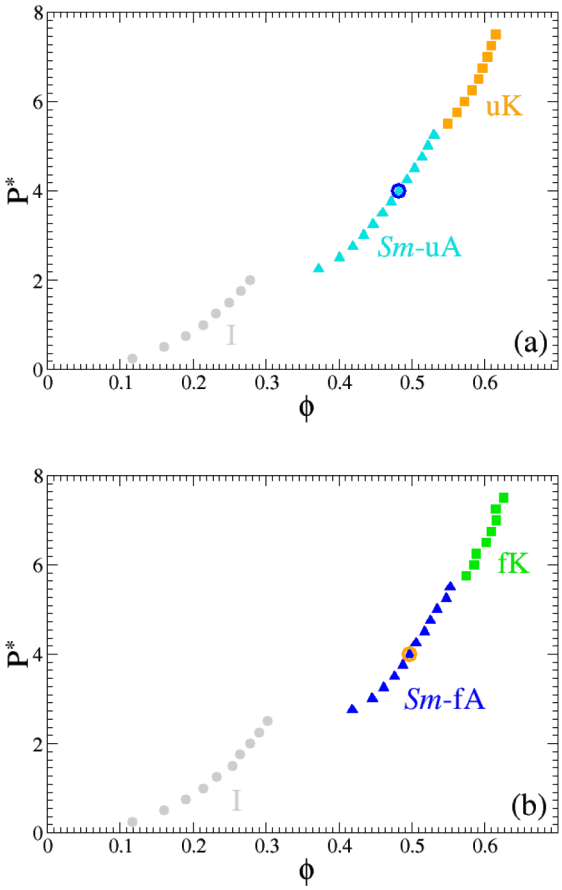

Fig. 3. Equation of state obtained from the two initial configurations shown in fig. 2 where gDNADs are either fully unfolded (a) or fully folded (b). In (a) the blue circle identifies the state point $\mathcal{U}$, while in (b) the orange circle identifies the state point $\mathcal{F}$.

one case all gDNADs are fully folded (labelled as fK, see fig. 2(b)) while in the other case they are fully unfolded (labelled as uK, see fig. 2(a)). During these simulations we equilibrated the system for at least 8×10^6 MC steps and we performed a production run lasting 9×10^6 MC steps.

The two equations of states obtained from these two series of MC simulations are shown in fig. 3, where the one in fig. 3(a) has been obtained from the initial configuration in fig. 2(a) and that in fig. 3(b) from the initial configuration in fig. 2(b). At high concentration the system can be found in a crystalline phase where the gDNADs are either unfolded (uK in fig. 3(a)) or folded (fK in fig. 3(b)). By reducing the pressure, *i.e.* lowering the concentration, the system exhibits a lamellar liquid-crystalline ordering. In the smectic phase of fig. 3(a) the majority of gDNADs are unfolded and this phase has been called *Sm*-uA, while in that of fig. 3(b) the majority of gDNADs are folded and this one has been called *Sm*-fA. The possibility of obtaining from simulations two different phases under the same thermodynamic conditions suggests that their free energy difference has to be small. Indeed, one of the two phases is metastable but being unable to undergo a transition to the more stable phase means that an high free energy barrier separates them. The latter point will be also discussed thoroughly later on.

A snapshot of the two different smectic phases obtained from our MC simulations for $P^* = 4$ is shown in fig. 4.

In fig. 5 we show the angle distribution $p(\theta)$, where θ is the angle formed by the two cylinders forming a gDNAD as shown in fig. 1. From figs. 4 and 5 it can be clearly seen that in the *Sm*-fA phase most of gDNADs are unfolded, while in the *Sm*-uA they are mostly unfolded, as also confirmed by the angle distribution shown in fig. 5.

2.3 Thermodynamic integration

To assess which of the two equations of state shown in figs. 3(a) and (b) is the one thermodynamically stable, we compute through thermodynamic integration the Gibbs free energy difference ΔG between two state points at the

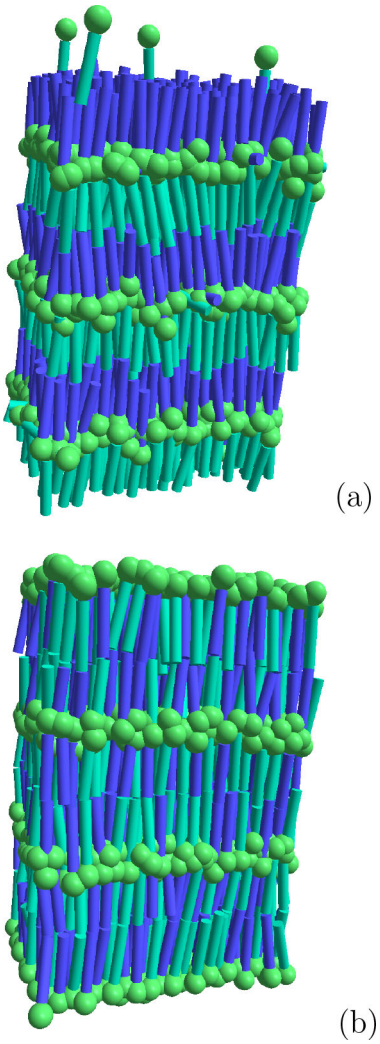

Fig. 4. Snapshots of (a) folded (Sm-fA) at $\phi = 0.497$ and (b) unfolded (Sm-uA) at $\phi = 0.482$ smectic phases.

same temperature and pressure. We chose the two state points $\mathcal{U}$ (see fig. 3(a)) and $\mathcal{F}$ (see fig. 3(b)) at $P^* = 4$ and $T^* = 0.125$, which correspond to the Sm-uA phase ($\phi = 0.482$) and to Sm-fA phase ($\phi = 0.497$), respectively and we build a suitable reversible path joining these two points in the phase diagram as shown in fig. 3 along which we evaluate the free energy change. Note that the chosen value of the reduced temperature T^* ensures that the interaction strength $\beta u_0 = 8$.

Along the reversible path, which we construct, both isothermal and Hamiltonian thermodynamic integration are needed [38,30]. Along an isotherm joining two states $\mathcal{A}$ and $\mathcal{B}$ the Gibbs free energy variation ΔG is

$$\Delta G = \int_{\mathcal{A}}^{\mathcal{B}} \frac{P(\rho)}{\rho^2}\,\mathrm{d}\rho + \frac{P(\mathcal{B})}{\rho(\mathcal{B})} - \frac{P(\mathcal{A})}{\rho(\mathcal{A})}\,. \tag{3}$$

In the case of Hamiltonian integration the Hamiltonian changes according to a parameter λ and goes from a state $\mathcal{A}$ ($\lambda = 0$) to a state $\mathcal{B}$ ($\lambda = 1$) and ΔG can be expressed

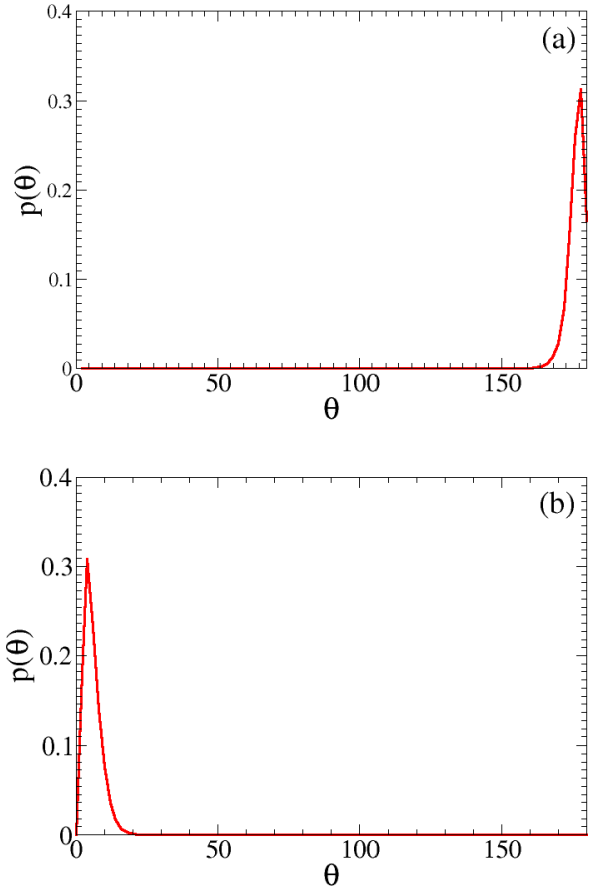

Fig. 5. Angle distribution for the unfolded and folded smectic state whose snapshots are shown in fig. 4.

as follows:

$$\Delta G = G(\mathcal{B}) - G(\mathcal{A}) = \int_{\mathcal{A}}^{\mathcal{B}} H_\lambda \mathrm{d}\lambda, \tag{4}$$

where

$$H_\lambda = \left\langle \frac{\mathrm{d}\mathcal{H}}{\mathrm{d}\lambda} \right\rangle_{N,P,T,\lambda}. \tag{5}$$

The Hamiltonian $\mathcal{H}_0$ of our system can be written as follows:

$$\mathcal{H}_0 = V_{HS} + \sum_{i<j} \sum_{a=1,2} \sum_{b=1,2} u_{SW}(\mathbf{r}_{ij}^{ab}) + \sum_i u_P(\mathbf{r}_i), \tag{6}$$

where V_{HS} is the hard-core potential acting between cylinders, $\mathbf{r}_{ij}^{ab}$ is the distance between the reversible interacting B sites belonging to cylinders i and j and $\mathbf{r}_i$ is the distance between two A sites belonging to the same gDNAD.

To build a suitable thermodynamic path connecting the latter two states we employed the following

4 correction terms to the original Hamiltonian $\mathcal{H}_0$:

$$\mathcal{H}^{(s)}(\lambda_s) = C_s \lambda_s u_0 \sum_i \cos\left(\frac{2\pi z_i N_p}{L_z}\right), \tag{7}$$

$$\mathcal{H}^{(n)}(\lambda_n) = C_n \lambda_n u_0 \sum_i \sum_{a=1,2} \left[1 - \cos(\mathbf{u}_{i,a} \cdot \hat{\mathbf{z}})\right], \tag{8}$$

$$\mathcal{H}^{(f)}(\lambda_f) = C_f \lambda_f u_0 \sum_i \mathbf{u}_{i,1} \cdot \mathbf{u}_{i,2}, \tag{9}$$

$$\mathcal{H}^{(p)}(\lambda_p) = -\lambda_p \sum_{i<j} \sum_{a=1,2} \sum_{b=1,2} u_{SW}(\mathbf{r}_{ij}^{ab}), \tag{10}$$

where we set $C_n = 1$, $C_s = 3$ and $C_f = 5$ and it is assumed that the nematic axis is along the z-axis. The constants C_γ with $\gamma = s, n, f$ have been chosen to ensure a sufficient strength of the associated external fields, *e.g.* when $\lambda_f = 1$ by setting $C_f = 5$ ensures that all gDNADs are perfectly folded despite thermal fluctuations. The correction $\mathcal{H}^{(s)}$ will induce, on varying λ_s, a transition to a smectic phase where B sites are placed onto layers perpendicular to the z-axis [30], while $\mathcal{H}^{(n)}$ will force the alignment of the cylinders along the z-axis on increasing the parameter λ_n [39]. The Hamiltonian correction $\mathcal{H}^{(f)}$ on increasing the parameter λ_f induces the folding of each gDNAD in the system and finally the correction $\mathcal{H}^{(p)}$ on changing λ_p from 0 to 1 completely switches off the attraction potential between A sites. By applying the above corrections to the original Hamiltonian $\mathcal{H}_0$ we induce the formation of an ideal folded smectic phase from state $\mathcal{F}$. From this state, we carry out MC simulations in the NPT ensemble from the pressure of the latter "ideal" (*i.e.*, with all particles almost perfectly aligned and layered) folded smectic phase down to very low pressures (and densities). The key idea of our procedure is that at low concentrations we can safely switch off all the fields, thus inducing the formation of an isotropic phase without encountering a first-order phase transition. After that, we turned back on all the corrections terms in eqs. (7)–(10), but, differently to what done before, the parameter λ_f is now varied from 0 to -1 in order to have gapped duplexes in unfolded configurations.

The detailed procedure to evaluate the Gibbs free energy difference between states $\mathcal{F}$ and $\mathcal{U}$ consists in splitting the whole integration path into the following intermediate steps, which are labelled as S_i with $i = 1, \ldots, 8$.

S_1: The smectic field in eq. (7) is gradually switched on and the system goes from state $\mathcal{F}$ to an ideal smectic state.

S_2: Hamiltonian correction terms in eqs. (8), (9) and (10) are gradually and simultaneously (*i.e.*, we vary $\lambda = \lambda_f = \lambda_n = \lambda_p$) switched on, so that eventually all gDNADs are perfectly folded and aligned.

S_3: The pressure is reduced until the system reaches very low concentrations.

S_4: All fields are gradually switched off thus obtaining an isotropic phase of gDNADs ($\mathcal{I}$).

S_5: All Hamiltonian corrections are gradually and simultaneously switched back on, where in the correction term $\mathcal{H}^{(f)}$, λ_f now changes from 0 to -1 to induce the formation of unfolded configurations.

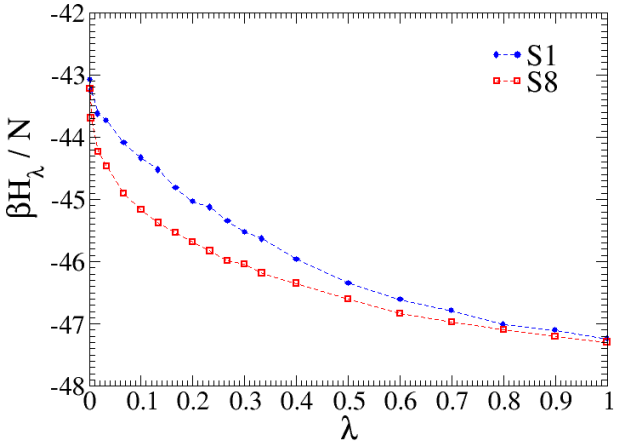

Fig. 6. The behavior of H_λ as a function of λ_s during steps S_1 (with gDNADs mostly folded) and S_8 (with gDNADs mostly unfolded) where the smectic field is switched on and off, respectively. All simulations are performed in the NPT ensemble at a pressure of $P^* = 4$.

S_6: The pressure is increased until the system reaches the same pressure of state $\mathcal{U}$.

S_7: Hamiltonian correction terms in eqs. (8), (9) and (10) are gradually and simultaneously (*i.e.*, we vary $\lambda = -\lambda_f = \lambda_n = \lambda_p$) switched off.

S_8: The smectic field (*i.e.*, the correction term in eq. (7)) is gradually switched off and eventually the system is in the state $\mathcal{U}$.

In all steps $S_1, \ldots, S_8$ we used NPT-MC simulations and each step S_i brings the system from state $\mathcal{P}_{i-1}$ to state $\mathcal{P}_i$ with $\mathcal{P}_0 = \mathcal{F}$, $\mathcal{P}_5 = \mathcal{I}$ and $\mathcal{P}_8 = \mathcal{U}$. Note that each state depends on temperature T, pressure P and the set of parameters $\{\lambda_\alpha \mid \alpha \in \{s, n, f, p\}\}$. According to the above steps the whole path along which thermodynamic integration is carried out can be expressed as follows:

$$\mathcal{F} \to \mathcal{P}_1 \to \mathcal{P}_2 \to \mathcal{P}_3 \to \mathcal{P}_4 \to \mathcal{P}_5 \to \mathcal{P}_6 \to \mathcal{P}_7 \to \mathcal{U}.$$

3 Results and discussion

In the following we will discuss the results from simulations for all the steps $\mathcal{S}_1, \ldots, \mathcal{S}_8$ into which we split the whole reversible path used for the thermodynamic integration. Figure 6 shows the behavior of H_λ as a function of λ_s when the smectic field is first switched on in step S_1 and then switched back off in S_8. The Gibbs free energy difference going along these two paths is $0.37 N k_B T$.

Figure 7 shows the behavior of H_λ as a function of λ when all fields, except the smectic one, are first switched on in step S_2 and then switched back off in step S_7. The change in the Gibbs free energy along these two paths is a bit smaller than the previous case, being equal to $-0.07 N k_B T$.

Finally, fig. 8 shows the behavior of the pressure P against ϕ along the isotherms which bring the system from state $\mathcal{P}_4$ to the low pressure state $\mathcal{I}$ ($= \mathcal{P}_5$) first and then

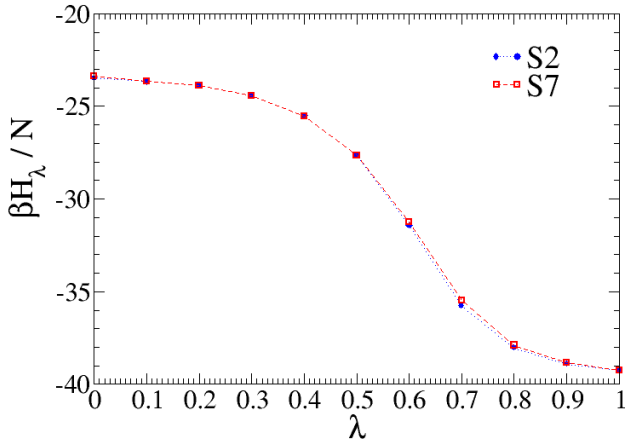

Fig. 7. H_λ as a function of λ in steps S_2 and S_7, where $\lambda = \lambda_n = \lambda_p = \lambda_f$ in step S_2 and $\lambda = \lambda_n = \lambda_p = -\lambda_f$ in step S_7. All simulations are performed in the NPT ensemble at $P^* = 4$.

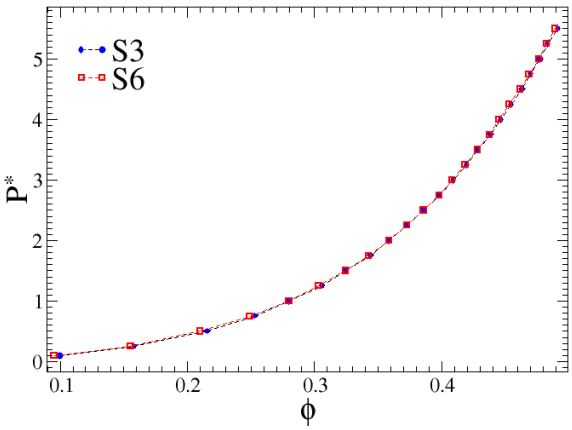

Fig. 8. Equation of state obtained in steps S_3 and S_6.

back to pressure $P^* = 4$ (state $\mathcal{P}_6$). The Gibbs free energy difference along these two paths is $0.25Nk_BT$.

Finally, fig. 9 shows the behavior of H_λ as a function of λ in steps S_4 and S_5, in this case the free energy variation along these two paths is $-0.30Nk_BT$. Note that this value is comparable in magnitude to that obtained in steps S_1 and S_8, since the range of values in the y-axis of fig. 9 is much larger than that in fig. 6.

We can conclude that the difference in Gibbs free energy going from state $\mathcal{F}$ to state $\mathcal{U}$ has comparable contributes from all integration steps. Overall the free energy difference between these two states is

$$\begin{aligned}\frac{\beta\Delta G}{N} &= \frac{\beta G(\mathcal{U})}{N} - \frac{\beta G(\mathcal{F})}{N} \\ &= \frac{\beta}{N}\left(\int_{\mathcal{F}}^{\mathcal{I}} H_\lambda \mathrm{d}\lambda + \int_{\mathcal{I}}^{\mathcal{U}} H_\lambda \mathrm{d}\lambda\right) \\ &= 0.25 \pm 0.05, \end{aligned} \tag{11}$$

where in each MC run the error has been carefully estimated using the method of block averages and then the overall error has been computed by considering error propagation which results from the integrals in eqs. (4) and (3) evaluated along the reversible path joining states $\mathcal{F}$ and $\mathcal{U}$.

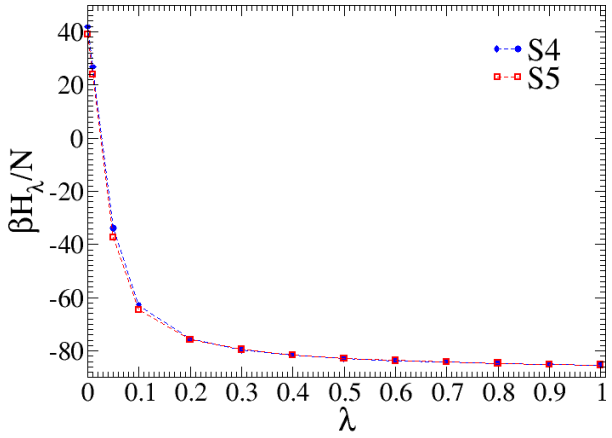

Fig. 9. H_λ during steps S_4 and S_5 of thermodynamic integration.

It can be seen that the difference is positive, *i.e.*:

$$G(\mathcal{F}) < G(\mathcal{U}) \tag{12}$$

therefore we conclude that the folded smectic state $\mathcal{F}$ is thermodynamically stable while the unfolded one $\mathcal{U}$ is metastable.

We note that the Gibbs free energy difference obtained from thermodynamic integration between states $\mathcal{F}$ and $\mathcal{U}$ is rather small [38]. For the sake of comparison, the very small free energy difference between the HCP (metastable) and the FCC (stable) crystals of hard spheres amounts to about $0.001Nk_BT$ [40–43] and the HCP phase can remain stable over a very long time [44]. Such a small value of the free energy difference between $\mathcal{F}$ and $\mathcal{U}$ states suggests that a large free energy barrier has to separate these two states. Within the classical nucleation theory the rate of nucleation $\mathcal{R}$ can be expressed as [45]

$$\mathcal{R} \propto \exp\left(-\frac{\Delta G^*}{k_BT}\right), \tag{13}$$

where

$$\Delta G^* \propto \frac{\gamma^3}{(\Delta\mu)^2}, \tag{14}$$

γ is the surface tension between the two phases $\mathcal{F}$ and $\mathcal{U}$ and $\Delta\mu = \Delta G/N$ is the difference between the chemical potential per particle of the two phases. Since in our case ΔG can be considered small, eq. (14) suggests that a large free energy barrier ΔG^* can be expected to separate these two states, and the nucleation rate in eq. (13) is likely to be small. Accordingly, if the $\mathcal{U}$ phase forms during a simulation, it will be unlikely for it to undergo a transition to the stable phase $\mathcal{F}$, as we observed in our simulations.

Now, we would like to estimate this free energy difference for any other value of the pressure at $T^* = 0.125$ where both phases can be obtained from simulations. To start, note that, if one has the equation of state for a given temperature T, the Gibbs free energy of a state

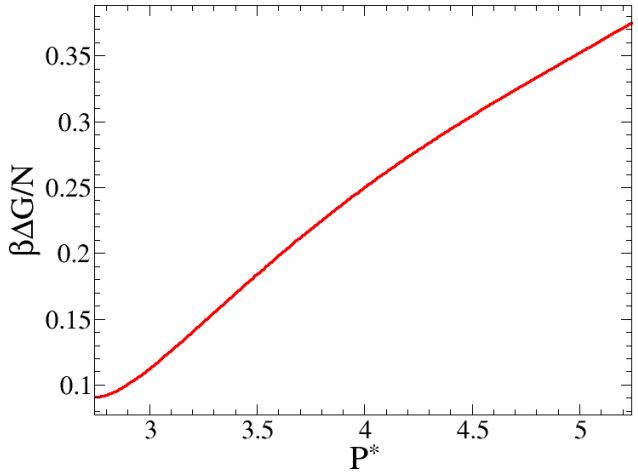

Fig. 10. Gibbs free energy difference between the *Sm*-uA and *Sm*-fA ΔG as a function of pressure for $T^* = 0.125$ obtained from eq. (16).

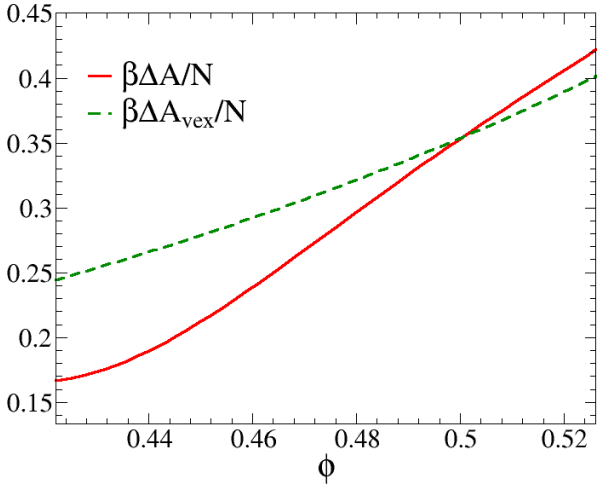

Fig. 11. Helmholtz free energy difference between the *Sm*-uA and *Sm*-fA as a function of pressure for $T^* = 0.125$. The solid line (ΔA) is the estimate obtained from G according to eq. (17), while the dashed line (ΔA_{ex}) is the free energy estimate based on excluded volume effects in eq. (18).

(P_n, T) can be obtained from that of a state (P_o, T) along an isotherm by using eq. (3), *i.e.*:

$$\frac{G(P_n,T)}{N} = \frac{G(P_o,T)}{N} + \int_{\rho_o}^{\rho_n} \frac{P(\rho)}{\rho^2}\,\mathrm{d}\rho + \frac{P_n}{\rho_n} - \frac{P_o}{\rho_o}\,, \quad (15)$$

where ρ_o and ρ_n are the number densities of the two states (P_o, T) and (P_n, T) respectively. From the latter equation the Gibbs free energy difference between the two smectic phases for any value of pressure P can be expressed as

$$\frac{\Delta G(P)}{N} = \frac{\Delta G(P_o)}{N} + \int_{\rho^u(P_o)}^{\rho^u(P)} \frac{P^u(\rho)}{\rho^2}\,\mathrm{d}\rho + \frac{P}{\rho^u(P)} - \frac{P_o}{\rho^u(P_o)} - \left[\int_{\rho^f(P_o)}^{\rho^f(P)} \frac{P^f(\rho)}{\rho^2}\,\mathrm{d}\rho + \frac{P}{\rho^f(P)} - \frac{P_o}{\rho^f(P_o)}\right], \quad (16)$$

where ρ^u and P^u are the density and pressure of the unfolded phase obtained from its equation of state and ρ^f and P^f are the same quantities but for the folded phase. To perform the integrals in eq. (16) we conveniently fit the equation of state to a quadratic polynomial. The resulting Gibbs free energy difference between the two smectic phases as a function of the pressure P^* is shown in fig. 10.

It can be seen that the free energy difference ΔG increases on increasing P. This behavior can be understood by considering that the folded configuration of gDNAD has a smaller excluded volume than the folded one and the resulting entropic difference scales as ρ^2.

Since the Helmholtz free energy is

$$A = G - PV \quad (17)$$

we can estimate from G its value at $T^* = 0.125$ on varying ρ as well as the free energy difference ΔA on changing the density. The result of these calculations is shown in fig. 11 for a range of concentration where both smectic phases can exist. As expected, one can conclude that at a given concentration, where both phases can be obtained from simulations, the *Sm*-fA phase is stable and the *Sm*-uA one is metastable.

Figure 11 shows also that free energy difference A increases on increasing concentration due to excluded volume effects, thus confirming what observed for ΔG on changing P.

The free energy $A_{vex}(\phi, T)$, accounting only for excluded volume effects, can be roughly expressed as follows [25]:

$$\frac{\beta A_{vex}}{N} = \frac{\beta A_{id}}{N} + \eta(\phi)\frac{\phi}{4v_{hc}}\overline{v}_{excl}, \quad (18)$$

where $\alpha \in \{f, u\}$, A_{id} is the ideal gas contribution, $\overline{v}_{excl}$ is the average excluded volume in the system and $\eta(\phi)$ is the Parsons-Lee factor, *i.e.*

$$\eta(\phi) = \frac{4 - 3\phi}{4(1-\phi)^2} \quad (19)$$

which accounts for higher-order terms in the virial expansion. Using the same method discussed in ref. [5], the excluded volumes in the smectic phase v^f_{excl} and v^u_{excl} of perfectly folded and unfolded conformations, respectively, can be calculated. It turns out that $v^f_{excl} = 12.3D^3$ and that $v^u_{excl} = 17.0D^3$. If we assume that in a *Sm*-uA state all gDNADs are fully unfolded while in a *Sm*-fA at the same concentration and temperature they are all fully folded, one has that $\overline{v}_{excl} = v^u_{excl}$ in the unfolded state and $\overline{v}_{excl} = v^u_{excl}$ in the unfolded state. Hence, an estimate of the free energy difference ΔA_{vex} between these two states can be provided, *i.e.*

$$\frac{\beta \Delta A_{vex}}{N} = \eta(\phi)\frac{\phi}{4v_{hc}}\left(v^u_{excl} - v^f_{excl}\right). \quad (20)$$

In fig. 11 we show ΔA_{vex} as a function of ϕ. Although we obtained eq. (18) by employing the Parsons-Lee approximation, which is known to be quite crude at larger

concentrations [25], the values of ΔA_{vex} are comparable with those obtained from eq. (17), thus suggesting that excluded volume effects contribute significantly to the stability of the *Sm*-fA phase. We also observe that ΔA_{vex} is expected to overestimate excluded volume effects, since in the *Sm*-fA (*Sm*-uA) only a fraction of gDNADs is fully folded (unfolded).

In addition, bond formation between gDNADs, which is neglected in eq. (18), must also contribute to the stability of the *Sm*-fA phase. The reason is twofold, on the one hand in the *Sm*-fA phase double bonding between two gDNADs stabilizes folded conformations. On the other hand, the aggregation of gDNADs induces the formation of semi-flexible chains, whose length distribution is expected to be exponential [25], thus ending up with a polydisperse set of chains in the system. As a consequence, the *Sm*-uA phase requires a higher entropic cost to form than the *Sm*-fA phase.

4 Conclusions

In this work we studied a simple coarse-grained model of DNA nanonunchakus which has been shown to exhibit both numerically and experimentally a smectic-A phase [5]. Depending upon the initial configuration used in the simulations, two different smectic-A phases can be obtained under identical thermodynamic conditions, where in one phase (*Sm*-fA) the particles are mostly folded and in the other (*Sm*-uA) they are mostly unfolded. In addition, we found that none of the two phases undergoes a transition to the other one during the simulation time span. Therefore, if one prepares the system in a crystalline state where all gDNADs are mostly folded or mostly unfolded, two different equations of state are obtained. By carrying out MC simulations we devised a protocol to estimate for a given temperature the free energy difference between any two smectic state points belonging to these two equations of states at the same concentration or pressure. In our approach we introduce suitable external potentials (which induce the formation of "ideal" smectic and nematic states) and interaction potentials (which force a specific gDNAD conformation and switch off their attraction), so that the system can go along a suitable path, joining the two states, whose free energy difference has to be estimated, without any intervening first-order phase transition. Our calculations provide evidence that the folded smectic phase is stable while the unfolded phase is metastable, thus confirming what proposed in ref. [5]. Due to hydrophobic attraction between blunt ends of gDNADs, unfolded conformations make smectic-type ordering unfavorable, since these interactions would induce the formation of a polydisperse set of linear semi-flexible aggregates, thus increasing the entropic cost of forming a smectic layering. Moreover, an almost fully folded conformation of gDNADs suppresses the formation of linear polydisperse aggregates, as the ones emerging from an unfolded conformation, and reduces the excluded volume between gDNADs, so that smectic layering becomes thermodynamically more convenient. The present numerical calculations based on MC simulations confirm this scenario and suggests that the larger driving force towards the formation of the *Sm*-fA phase is given by excluded volume effects. The inability of the system to spontaneously undergo a transition from the *Sm*-fA to *Sm*-uA phase is possibly ascribed to their small free energy difference.

The numerical protocol based on the thermodynamic integration proposed here is rather general and it can be employed to assess the thermodynamic stability of any thermodynamic phase competing with metastable phases of similar free energy. When different phases obtained under the same thermodynamic conditions have a small free energy difference, by carrying out computer simulations, that start from different initial conditions, all of such phases can be obtained in principle. Under these circumstances, it is required to study the thermodynamic stability of these phases through an affordable approach, since a spontaneous transition to the most stable phase is unlikely to occur within the limited simulation time span. Our protocol overcomes this difficulty by allowing the direct calculation of the free energy difference between any pair of phases of similar free energy.

We can envisage many applications of our method, for example by changing the length of the spacer or of the two sticks which form each gDNADs new mesophases could emerge and possibly it will be required to verify their thermodynamic stability. Another system where our protocol could be applied can be provided by clathrate hydrates [46,47]. Hydrates clathrates are crystalline compounds composed of water (host) molecules and gas (guest) atoms or molecules [48,49]. Clathrate hydrates of methane are present on earth in permafrost areas and ocean shelves. They are believed to be the dominant methane-bearing phase in the nebula from which the outer planets and satellites are formed, hence their properties are also needed to model astronomical bodies in the outer solar system [50]. The interest in these materials stems also from the risk of their formation in oil and gas pipelines and their potential application as gas transportation media in soft conditions, *i.e.* close to atmospheric pressure and room temperature. Clathrate hydrates can be commonly found in two different structures, which are called type I (sI) and type II (sII) [51]. The unit cell of clathrate sI contains two small dodecahedral water cages and six bigger tetrakaidecahedral cages, while the unit cell of sII contains 16 5^{12} cages and eight large hexadecahedral cages (for a graphical representation of these two structures see fig. 1 in ref. [48]). The two structures sI and sII are known to have a small free energy difference [52], hence it would be valuable to have a convenient method to assess their thermodynamic stability over a large range of pressures at a given temperature. Our protocol can be in fact used to calculate the Gibbs free energy difference of these two structures over a range of pressures and temperatures, where they can both remain stable in computer simulations.

In conclusion, we are confident that beyond the present case study our approach could be a valuable tool to study the thermodynamic stability of phases, which can be obtained from computer simulations.

Authors warmly thank Prof. Francesco Sciortino for useful discussions.

Author contribution statement

All the authors were involved in the preparation of the manuscript. All the authors have read and approved the final manuscript.

References

1. M. Nakata, G. Zanchetta, B.D. Chapman, C.D. Jones, J.O. Cross, R. Pindak, T. Bellini, N.A. Clark, Science **318**, 1276 (2007).
2. G. Zanchetta, Liq. Cryst. Today **18**, 40 (2009).
3. G. Zanchetta, F. Giavazzi, M. Nakata, M. Buscaglia, R. Cerbino, N.A. Clark, T. Bellini, Proc. Natl. Acad. Sci. U.S.A. **107**, 17497 (2010).
4. C. Maffeo, B. Luan, A. Aksimentiev, Nucl. Acids Res. **40**, 3812 (2012).
5. M. Salamonczyk, J. Zhang, G. Portale, C. Zhu, E. Kentzinger, J.T. Gleeson, A. Jakli, C. De Michele, J.K.G. Dhont, S. Sprunt *et al.*., Nat. Commun. **7**, 13358 (2016).
6. G.P. Smith, T.P. Fraccia, M. Todisco, G. Zanchetta, C. Zhu, E. Hayden, T. Bellini, N.A. Clark, Proc. Natl. Acad. Sci. U.S.A. **115**, E7658 (2018).
7. S.D. Leo, M. Todisco, T. Bellini, T.P. Fraccia, Liq. Cryst. **45**, 2306 (2018).
8. T.P. Fraccia, G.P. Smith, N.A. Clark, T. Bellini, Crystals **8**, 5 (2017).
9. S. Saurabh, Y. Lansac, Y.H. Jang, M.A. Glaser, N.A. Clark, P.K. Maiti, Phys. Rev. E **95**, 032702 (2017).
10. R. Cortini, X. Cheng, J.C. Smith, J. Phys.: Condens. Matter **29**, 084002 (2017).
11. C. Robinson, Tetrahedron **13**, 219 (1961).
12. R.L. Rill, Proc. Natl. Acad. Sci. U.S.A. **83**, 342 (1986).
13. R. Brandes, D.R. Kearns, Biochemistry **25**, 5890 (1986).
14. D. Durand, J. Doucet, F. Livolant, J. Phys. II **2**, 1769 (1992).
15. F. Livolant, A.M. Levelut, J. Doucet, J.P. Benoit, Nature **339**, 724 (1989).
16. T.M. Alam, G. Drobny, J. Chem. Phys. **92**, 6840 (1990).
17. F. Livolant, A. Leforestier, Prog. Polym. Sci. **21**, 1115 (1996).
18. J. Pelta, D. Durand, J. Doucet, F. Livolant, Biophys. J. **71**, 48 (1996).
19. K. Merchant, R.L. Rill, Biophys. J. **73**, 3154 (1997).
20. H.H. Strey, R. Podgornik, D.C. Rau, V.A. Parsegian, Curr. Opin. Struct. Biol. **8**, 309 (1998).
21. R. Podgornik, H.H. Strey, V.A. Parsegian, Curr. Opin. Colloid Interface Sci. **3**, 534 (1998).
22. F. Tombolato, A. Ferrarini, J. Chem. Phys. **122**, 054908 (2005).
23. T.E. Strzelecka, M.W. Davidson, R.L. Rill, Nature **331**, 457 (1988).
24. F. Livolant, Y. Bouligand, J. Phys. (Paris) **47**, 1813 (1986).
25. C. De Michele, T. Bellini, F. Sciortino, Macromolecules **45**, 1090 (2012).
26. C. De Michele, L. Rovigatti, T. Bellini, F. Sciortino, Soft Matter **8**, 8388 (2012).
27. C. De Michele, G. Zanchetta, T. Bellini, E. Frezza, A. Ferrarini, ACS Macro Lett. **5**, 208 (2016).
28. K.T. Nguyen, A. Battisti, D. Ancora, F. Sciortino, C. De Michele, Soft Matter **11**, 2934 (2015).
29. M. Schoen, A.J. Haslam, G. Jackson, Langmuir **33**, 11345 (2017).
30. P. Bolhuis, D. Frenkel, J. Chem. Phys. **106**, 666 (1997).
31. A. Haji-Akbari, M. Engel, S.C. Glotzer, J. Chem. Phys. **135**, 194101 (2011).
32. G. Navascués, E. Velasco, Phys. Rev. E **95**, 032140 (2017).
33. B. Cheng, M. Ceriotti, Phys. Rev. B **97**, 054102 (2018).
34. M. Radu, P. Pfleiderer, T. Schilling, J. Chem. Phys. **131**, 164513 (2009).
35. M. Müller, K.C. Daoulas, J. Chem. Phys. **128**, 024903 (2008).
36. C. Greco, Y. Jiang, J.Z.Y. Chen, K. Kremer, K.C. Daoulas, J. Chem. Phys. **145**, 184901 (2016).
37. K.T. Nguyen, F. Sciortino, C. De Michele, Langmuir **30**, 4814 (2014).
38. C. Vega, E. Sanz, J.L.F. Abascal, E.G. Noya, J. Phys.: Condens. Matter **20**, 153101 (2008).
39. P.A. O'Brien, M.P. Allen, D.L. Cheung, M. Dennison, A. Masters, Phys. Rev. E **78**, 051705 (2008).
40. L.V. Woodcock, Nature **385**, 141 (1997).
41. P.G. Bolhuis, D. Frenkel, S.C. Mau, D.A. Huse, Nature **388**, 235 (1997).
42. K. Kendall, C. Stainton, F. van Swol, L.V. Woodcock, Int. J. Thermophys. **23**, 175 (2002).
43. E.G. Noya, N.G. Almarza, Mol. Phys. **113**, 1061 (2015).
44. I.P. Dolbnya, A.V. Petukhov, D.G.A.L. Aarts, G.J. Vroege, H.N.W. Lekkerkerker, Europhys. Lett. **72**, 962 (2005).
45. R. Zandi, P. van der Schoot, D. Reguera, W. Kegel, H. Reiss, Biophys. J. **90**, 1939 (2006).
46. C.A. Koh, A.K. Sum, E.D. Sloan, J. Appl. Phys. **106**, 061101 (2009).
47. D.K. Staykova, W.F. Kuhs, A.N. Salamatin, T. Hansen, J. Phys. Chem. B **107**, 10299 (2003).
48. U. Ranieri, M.M. Koza, W.F. Kuhs, S. Klotz, A. Falenty, P. Gillet, L.E. Bove, Nat. Commun. **8**, 1076 (2017).
49. S. Schaack, U. Ranieri, P. Depondt, R. Gaal, W.F. Kuhs, A. Falenty, P. Gillet, F. Finocchi, L.E. Bove, J. Phys. Chem. C **122**, 11159 (2018).
50. L. Bove, U. Ranieri, Philos. Trans. R. Soc. A **377**, 20180262 (2019).
51. E. Sloan, C. Koh, *Clathrate Hydrates of Natural Gases*, 3rd edition (CRC Press, Boca Raton, Florida, 2008).
52. L.C. Jacobson, V. Molinero, J. Am. Chem. Soc. **133**, 6458 (2011).