SAPIENZA
UNIVERSITÀ DI ROMA

Facoltà di Scienze Matematiche Fisiche e Naturali

Tesi di Dottorato in Scienza dei Materiali
XXV Ciclo

Role of the anisotropy in the interactions between nano– and micro–sized particles

Relatore
Prof. Francesco Sciortino

Candidato
Dr. Lorenzo Rovigatti

Contents

Introduction

The present Thesis focuses on the thermodynamic and dynamic behaviour of anisotropically interacting colloids by means of theoretical and numerical techniques.

Colloidal suspensions, i.e. micro– and nano–sized particles dispersed in a continuous phase, are a topic of great interest in several fields, including material science, soft matter and biophysics [1, 2]. Common in everyday life in the form of soap, milk, cream, etc., colloids have been used for decades as models for atomic and molecular systems, since both classes of systems share many features like critical phenomena, crystallisation and glass transition.

Experimental investigation of colloidal systems is made easier by the large size of colloids, which makes it possible to employ visible light as an experimental probe to investigate these systems. Moreover, since the mass of the particles controls the timescales of the dynamics, relaxation times of colloidal suspensions, ranging from seconds to years, orders of magnitude larger than their atomic counterparts, are more easily experimentally accessible. For example, a colloidal particle with diameter 1 μm in water and at ambient temperature diffuses a distance equal to its diameter in $\approx$ 0.38 s [3]. By exploiting this intrinsic slowness, with respect to molecular liquids, present day experimental techniques make it possible to follow in time trajectories of ensembles of particles with tools like confocal microscopy, thus effectively allowing to reconstruct the whole phase space trajectory of the system. In addition, it is also possible to manipulate single and multiple objects using techniques like optical tweezers, magnetic tweezers and atomic force microscopy [4, 5]. With single-molecule force spectroscopy one can arrange particles in ordered structures or measure properties like stiffness or mechanical responses (as in pulling experiments on RNA and DNA strands of particles and aggregates [6, 7]).

A remarkable difference between the molecular and the colloidal world is that in the former the interactions between the basic constituents are fixed by

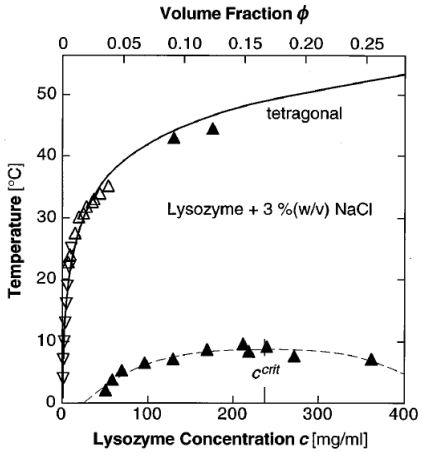

Figure 1: (Reproduced from Ref. [14]). Phase diagram for an aqueous lysozyme solution. The gas-liquid-like phase separation is metastable with respect to crystallisation.

nature, while in the latter the effective potential between two particles can be controlled by accurately designing and synthesizing the building blocks [8] or tuned by changing the properties of the solvent. This can be exploited, for example, to realize the simplest case of a stable colloidal dispersion: a hard-sphere system. Hard-spheres are particles that do not overlap and are non-interacting for separations larger than their diameters. Such a system, driven only by entropy, has been shown to exhibit crystallisation [9] and, if polydispersity is sufficiently high, a glass transition [10]. The addition of polymers in solution with a radius of gyration r_g much smaller than the colloids size leads to an effective attractive isotropic interaction between colloidal particles, whose range is controlled by r_g [11, 12]. This purely entropic attraction, the so-called *depletion* interaction, arises from an increasing of the free volume accessible to the polymers when two colloids are separated by a distance $r \lesssim r_g$. One can tune both the attraction range (changing the size of the polymer) and strength (changing the concentration of the polymer), with the latter playing the role of the inverse of the temperature. As it is known from classic results, the presence of an attractive term in the interaction potential leads to the appearance, at low temperatures, of a region of the phase diagram where a low-density, gas-like phase coexists with a high-density, liquid-like phase [13].

Differently from molecular systems, the coexisting region is shifted at lower temperatures and it is metastable with respect to the crystallisation of the supercritical fluid phase. Globular proteins often exhibit this behaviour. For

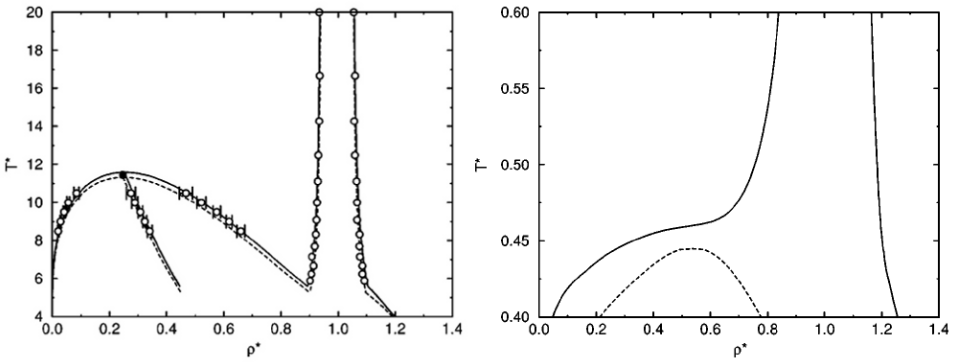

Figure 2: (Reproduced from Ref. [15]). Phase diagrams of systems with an isotropic attraction of the type $u(i,j) = -\epsilon\, r_{ij}^{-(3+\sigma)}$. The smaller σ the longer the range of the attraction. Left: $\sigma = 0.1$: the gas–liquid and fluid–solid coexistence regions are well separated. Right: $\sigma = 5$, the gas–liquid coexisting becomes metastable (dashed curve) with respect to crystallisation.

example, the low-concentration phase diagram of lysozyme is shown in Fig. 1. Fig. 2 shows that varying the range of the potential shifts the gas-liquid phase coexisting region with respect to crystallisation.

This behaviour, which does not occur in simple liquids, is due to the short-range nature of the attraction, whose range does not exceed, and often is considerably smaller than, the particle size [15,16]. Another important consequence of such short-range character of the interaction is the Noro–Frenkel law of corresponding states: gas–liquid coexisting curves of systems interacting through pair potentials with different functional forms can be rescaled on top of each other if the second virial coefficient B_2^* and the reduced density ρ^* are used in place of T and ρ [17,18].

Not only the thermodynamics, but also the dynamics of colloidal suspensions can exhibit fascinating and unexpected phenomena which have no correspondence in atomic and molecular systems. Dynamical arrest and glass transitions in molecular systems have direct counterparts in the colloidal world, but the latter displays a richer phenomenology due the larger variety of possible interactions. For example, it has been experimentally, numerically and theoretically shown that adding a very short attraction to a hard-sphere system can lead to the appearance of a glass at low temperature and intermediate-high densities which is driven by attraction rather than by repulsion [19–21]. Depending on the interactions, the two qualitatively different glasses can be separated by both a glass-glass transition and a pocket of liquid states. These equilibrium states can be accessed by preparing the sample at the right volume fraction and either

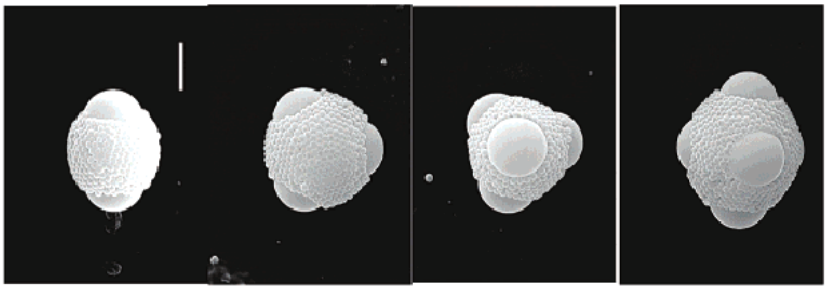

Figure 3: (Reproduced from Ref. [23]). Experimental particles realised from bidisperse colloids in water droplets.

by heating the attractive glass or by cooling the repulsive glass, providing an example of reentrant melting. Since glass transition *locii* can be tuned by changing the details of the interaction, this phenomenon can be exploited by material scientists to design and produce new technologically relevant materials whose mechanical properties, e.g. stiffness, can be varied by order of magnitudes by slightly changing the external parameters [20, 22].

In the last decade many new sophisticated techniques for particle synthesis have been developed and refined. These recent advances allow for the creation of an incredible variety of non-spherically, i.e. anisotropically, interacting building blocks. The anisotropy can arise from shape, surface patterning, form of the interactions or a combination thereof. Examples are colloidal cubes [24, 25], Janus particles [26, 27], triblock Janus particles [28], patchy particles [23, 29, 30], magnetic spheres [31] and many others. Fig. 3 shows an experimental realisation of patchy colloids. In order to classify and rationalise the anisotropy space, it has been recently suggested that the anisotropy of particles can be projected on different anisotropy "dimensions" [8], with some examples provided in Fig. 4. For instance, a hard ellipsoid with attractive spots on its surface can be described in terms of aspect ratio, number, and width and range of the spots. As a result, different kind of such ellipsoids could be considered as points of a three-dimensional anisotropy space whose axes represent the anisotropy dimensions A, B and D of Fig. 4.

The recent blossoming of experimental, theoretical and numerical studies and research on the role of the anisotropy has highlighted the richness of phenomena that these systems exhibit. Relevant examples for the present Thesis are valence-limited building blocks, i.e colloids with a maximum number of bound neighbours, and non-spherical particles with an aspect ratio, i.e. the ratio of the width of a particle to its height, significantly different from 1.

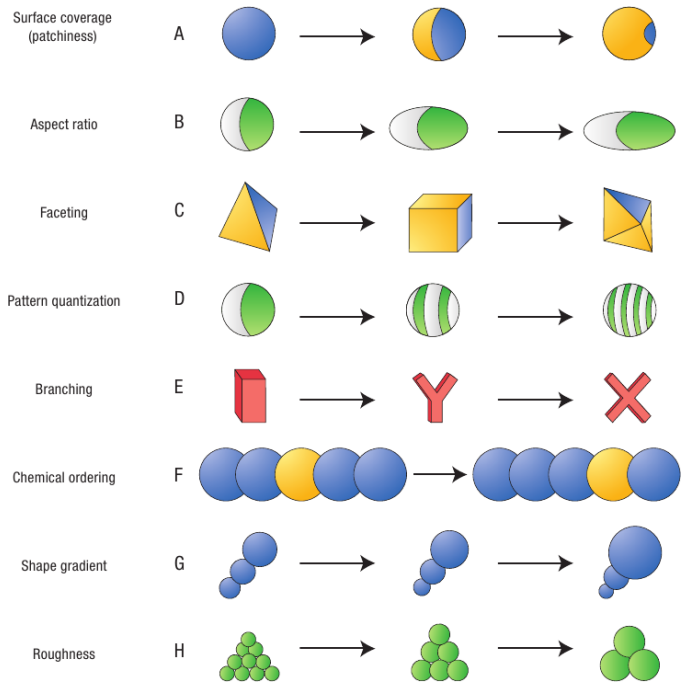

Figure 4: (Reproduced from Ref. [8]). Several examples of anisotropy "dimensions". The attribute corresponding to the given anisotropy axis is varied left to right.

The simplest example of valence-limited colloids is given by the so-called *patchy* particles: colloids decorated with attractive spots (patches) on the surface [29, 30, 32]. If the width and the range of the patches are chosen in such a way that each patch can form no more than one bond, then the total number of bound first neighbours per particle M can not exceed the number of patches. For particles interacting through short-ranged isotropic potentials, $M \approx 12$. It has been shown that changing the valence M has dramatic effects, both qualitative and quantitative, on the dynamic and thermodynamic properties of such systems. At high densities patchy colloids can self-assemble into a large variety of crystal structures, depending on valence, geometry and external parameters [33–36]. We will mostly focus on low-density systems. Since percolation is a prerequisite for the liquid state, in the limit $M \to 2$, when no branching is possible, the gas–liquid transition disappears and, at low temperatures, the system undergoes an equilibrium polymerization process [37]. Gradually lowering the valence from the isotropic case to the $M \to 2$ limit makes the gas–liquid phase coexisting region shrink towards lower temperatures and densities [38, 39]. This opens up a region of the phase diagram where liquid-like states can be explored in equilibrium by lowering the temperature, without encountering

phase separation. Such states, called "empty liquids", are self-assembled, low-density percolating networks whose bond lifetime, and hence the most relevant microscopic timescale, can be controlled by varying the temperature [3, 40]. Upon lowering T the system undergoes a dynamical arrest in a continuous way, providing a route for the generation of ideal, i.e. thermoreversible, gels [41, 42].

The second class of systems pertinent to the present work comprises anisotropically shaped particles that, depending on the aspect ratio and the values of the external parameters, can exhibit liquid crystal phases which may display orientational long-range order [43]. Nematic, in which there is no translational order, smectic, in which particles are ordered in layers and thus exhibit translational order in one dimension, and columnar phases, in which particles self-assemble into cylindrical aggregates which can in turn become nematic or form two-dimensional lattices, do not exist in isotropic systems, since the anisotropy in shape is a prerequisite for the breaking of the orientational symmetry. Liquid crystals, discovered at the end of the 19th century [44], have been thoroughly investigated for decades [45–48], leading to technological breakthroughs like LCD displays. Recently it has been suggested that liquid crystal phases occurring in dense solutions of short DNA double strands could have played a role in the prebiotic chemical generation of complementary H-bonded molecular assemblies [49].

The main goal of the present Thesis is to study the structural, thermodynamic and, to a lesser extent, dynamic properties of systems interacting through anisotropic potentials at low densities and temperatures. In particular, we focus on the low-density phase behaviour of valence-limited systems. We use a variegated approach, comprising state-of-the-art Monte Carlo and Molecular Dynamics techniques, to analyse and shed some light on the effect of the anisotropy on the phase diagram and on the dynamics of such systems.

As the effect of the valence on the phase diagram plays a major role in the models investigated throughout this Thesis, each Chapter is devoted to the study of the dynamics and thermodynamics of systems having a fixed or effective maximum valence M.

In the last years a lot of effort has been devoted to the study of end-to-end stacking interactions between different strands of nucleic acids, which play an important role in both physical and biological applications of DNA and RNA [50, 51]. In Chapter 1, building on the experimental work of Bellini *et al.*, we make use of a theoretical framework recently developed [52] to tackle the problem of the isotropic–nematic phase coexistence in solutions of short DNA

duplexes (DNADs). We compare the parameter-free theoretical predictions with results from large scale numerical simulations on GPUs of a coarse-grained realistic model and find a good quantitative agreement at low concentrations. We then predict the phase boundaries for different DNAD lengths and compare the results with experimental findings [53].

In Chapter 2 we investigate the structural and thermodynamic properties of systems having $M = 2$, that is systems that undergo an extensive formation of linear structures as temperature is lowered. We focus on bi-functional patchy particles whose interaction details are chosen to qualitatively mimic the behaviour of the low-density, low-temperature dipolar hard sphere (DHS) model by analysing the outcomes of the simulations carried out in Chapter 3. In particular, we are interested in the interplay between chains and rings in equilibrium polymerization processes in a region of the phase diagram where the formation of the latter is favoured. The very good quantitative agreement found by comparing numerical results with theoretical, parameter-free predictions calls for an extension of the theory with the inclusion of branching, in order to understand how the presence of rings affects the phase separation [54].

Chapter 3 is devoted to the investigation of the phase behaviour of dipolar fluids, i.e. systems interacting mainly through dipole-dipole potentials. For spheres, the lowest-energy configuration is the nose-to-tail contact geometry, and hence the ground state is an infinite chain or ring like in regular $M = 2$ systems. For finite temperatures, on the other hand, thermal fluctuations allow for the appearance of defects like dangling ends and chain branching which, in the language of this Thesis, makes for a temperature-dependent valence. This general mechanism, under some specific conditions, can lead to a very peculiar phase separation, driven by a balance between these *topological* defects rather than by the energy/entropy competition usually responsible for regular gas–liquid phase transitions [55]. This topological phase transition has been recently observed in a model system of patchy particles [56,57] but it is unclear whether such mechanism still holds in dipolar fluids in general and in the DHS model in particular. We focus on the DHS model, whose phase behaviour at low densities and temperatures has been studied for decades [58–60] but still remains largely unknown. In particular, we look for the gas–liquid critical point by means of state-of-the-art Monte Carlo simulations in a region where it has long been thought to be. We find no evidence of a phase transition and we speculate that this is due to an abundance of rings, providing a remarkable example of phase separation suppressed by self-assembly [61,62].

In Chapter 4 we study the dynamics of tetravalent patchy particles in the optimal network density region. For this fixed value of density the system is able to form a fully connected random network, i.e. an ideal gel [63, 64]. Indeed, as the temperature is lowered, a percolating network forms and the dynamics slows down. Although the observed dynamical arrest is different from the glass case, where excluded volume interactions are dominant, the decay of the self– and collective correlation functions of the resulting fluid bears similarities with that observed in glassy systems. Remarkably, comparing the characteristic decay times of density-density correlation functions with the average bond life, we find that only at very low T the decay of the density fluctuations requires the breakage of bonds [65].

In Chapter 5 we introduce DNA as a building block that can be used to rationally design novel, self-assembling materials with tunable properties. In this Chapter, we study the phase behaviour of four-armed DNA constructs at low densities. We use the coarse-grained, realistic DNA model employed in Chapter 1 and state-of-the-art simulation techniques, as presented in Chapter 6, to investigate systems composed of thousands of nucleotides undergoing a two-step self-assembling process and we quantitatively compare the outcome with experimental results obtained for a very similar system.

In Chapter 6 we introduce Graphics Processing Units (GPUs) as valuable tools for present day numerical investigations. We outline both the architecture of NVIDIA GPUs and NVIDIA CUDA, the software layer built on top of the hardware required to program these devices. We then present the techniques employed to write an efficient, general Molecular Dynamics code and compare its performances with a regular CPU code. The observed performance boost allows us to tackle the analysis of the dynamics and thermodynamics of very large systems without having to resort to massive CPU clusters (see Chapters 1, 4 and 5).

Chapter 1

Valence two: self-assembly of short DNA duplexes

1.1 Introduction

In this Chapter and in the following one we will focus on systems with valence two. Upon lowering the temperature, polymerisation takes place and the basic constituents self-assemble into linear aggregates such as chains or rings.

Self-assembly is the spontaneous formation through free energy minimization of reversible aggregates of basic building blocks. The size of the aggregating units, e.g. simple molecules, macromolecules or colloidal particles, can vary from a few angströms to microns, thus making self-assembly ubiquitous in nature and of interest in several fields, including material science, soft matter and biophysics [1, 2, 23, 66, 67]. Through self-assembly it is possible to design new materials whose physical properties are controlled by tuning the interactions of the individual building blocks [68–71].

A relevant self-assembly process is the formation of filamentous aggregates (i.e. linear chains) induced by the anisotropy of attractive interactions. Examples are provided by micellar systems [72–74], formation of fibers and fibrils [75–78], solutions of long duplex B-form DNA composed of 10^2 to 10^6 base pairs [79–82], filamentous viruses [83–87], chromonic liquid crystals [88] as well as inorganic nanoparticles [89].

If linear aggregates possess sufficient rigidity, the system may exhibit liquid crystal (LC) phases (e.g. nematic or columnar) above a critical concentration. In the present study we focus on the self-assembly of short (i.e. 6 to 20 base

pairs) DNA duplexes (DNADs) [49,90,91] in which coaxial stacking interactions between the blunt ends of the DNADs favor their aggregation into weakly bonded chains. Such a reversible physical polymerization is enough to promote the mutual alignment of these chains and the formation of macroscopically orientationally ordered nematic LC phases. At present, stacking is understood in terms of hydrophobic forces acting between the flat hydrocarbon surfaces provided by the paired nucleobases at the duplex terminals, although the debate on the physical origin of such interaction is still active and lively [51,92]. In this respect, the self-assembly of DNA duplexes provides a suitable way to access and quantify hydrophobic coaxial stacking interactions.

In order to extract quantitative informations from DNA-DNA coaxial stacking experiments, reliable computational models and theoretical frameworks are needed. Recent theoretical approaches have focused on the isotropic-nematic (I-N) transition in self-assembling systems [93,94], building on previous work on rigid and semi-flexible polymers [95–103]. In a recent publication [52], the reversible physical polymerization and collective ordering of DNA duplexes was investigated by modelling them as super-quadrics with quasi-cylindrical shape [104] with two reactive sites [105, 106] on their bases. The theoretical framework presented therein, built on Wertheim [107–109] and Onsager [45] theories, was shown to be able to properly account for the association process.

Here, we employ that theoretical framework to study the physical properties of a realistic coarse-grained model of DNA recently proposed by Ouldridge *et al.* [110], where nucleotides are modelled as rigid bodies interacting with site-site potentials. Following Ref. [52], we compute the inputs required by the theory, i.e. the stacking free energy and the DNAD excluded volume, for the Ouldridge *et al.* model [110]. Subsequently we predict the polymerization extent in the isotropic phase as well as the isotropic-nematic phase boundaries.

To validate the theoretical predictions, we perform large-scale molecular dynamics (MD) simulations in the NVT ensemble of a bulk system comprising 9600 nucleotides, a study made possible by the computational power of modern Graphical Processing Units (GPUs). The parameter-free theoretical predictions provide an accurate description of the simulation results in the isotropic phase, supporting the theoretical approach and its application in the comparison with experimental results.

The Chapter is organized as follows. Section 1.2 provides details of the coarse-grained model of DNADs and of the MD computer simulations, Section 1.3 presents a summary of the theory, developed by Dr. C. De Michele.

Section 1.4 describes the protocols implemented to evaluate the input parameters required by the theory via MC integrations for two DNADs. We also discuss some geometrical properties of the bonded dimer configurations. We then compare the theoretical predictions with simulation and experimental results. Finally, in Section 1.5 we discuss estimates for the stacking free energy and present our conclusions.

1.2 Methods

1.2.1 Model

We implement a coarse-grained model for DNA recently developed by Ouldridge *et al.* [110, 111]. For a complete description we refer the reader to Refs. [110] and [112].

The model represents DNA as a string of nucleotides, where each nucleotide (sugar, phosphate and base group) is a rigid body with interaction sites for backbone, stacking and hydrogen-bonding interactions. The potential energy of the system is

$$\begin{aligned} V_0 &= \sum_{\langle ij \rangle} \left(V_{\text{b.b.}} + V_{\text{stack}} + V'_{\text{exc}} \right) + \\ &+ \sum_{i,j \notin \langle ij \rangle} \left(V_{\text{HB}} + V_{\text{cr.st.}} + V_{\text{exc}} + V_{\text{cx.st.}} \right), \end{aligned} \tag{1.1}$$

where the first sum is taken over all nucleotides that are nearest neighbours on the same strand and the second sum comprises all remaining pairs. The interactions between nucleotides are schematically shown in the Fig. 1.1, and the explicit forms can be found in Refs. [110] and [112]. The hydrogen bonding (V_{HB}), cross stacking ($V_{\text{cr.st.}}$), coaxial stacking ($V_{\text{cx.st.}}$) and stacking interactions (V_{stack}) explicitly depend on the relative orientations of the nucleotides as well as on the distance between interaction sites. The backbone potential $V_{\text{b.b.}}$ is an isotropic spring that imposes a finite maximum distance between neighbours, mimicking the covalent bonds along the strand. The coaxial stacking term, not shown in the Fig. 1.1, is designed to capture stacking interactions between non-neighbouring bases, usually on different strands. All interaction sites also have isotropic excluded volume interactions V_{exc} or V'_{exc}.

The coaxial-stacking interaction, particularly relevant for the present topic, acts between any two non-bonded nucleotides and is responsible for the duplex-

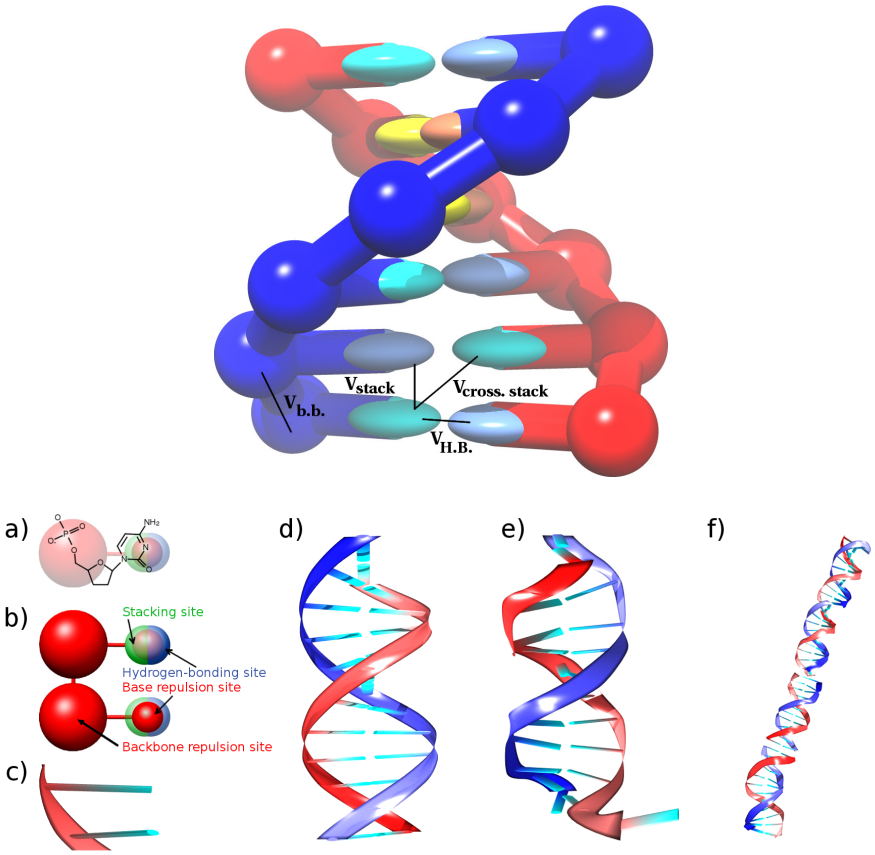

Figure 1.1: Top: interactions between nucleotides in the coarse-grained DNA model for two strands in a duplex. All nucleotides also interact with a repulsive excluded volume interactions. The coaxial stacking interaction is not shown. Bottom: (a) Schematic representation of the coarse graining of the model for a single nucleotide. (b) Model interaction sites. For the sake of clarity, stacking and hydrogen-bonding sites are highlighted on one nucleotide and the base repulsion site on the other. For visualization reasons, in the following strands will be shown as ribbons and bases as extended plates as depicted in (c). (d) A 12 base-pairs DNAD in the minimum energy configuration. (e) An equilibrium configuration of the same object at $T = 300$ K. The nucleotides at the bottom are not bonded, the so-called fraying effect. (f) A chain of length $N_b = 48$ extracted from a simulation at $c = 241$ mg/ml and $T = 270K$.

duplex bonding. It has been parametrized [112] to reproduce experimental data which quantify the stacking interactions by observing the difference in the relative mobility of a double strand where one of the two strands has been nicked with respect to intact DNA [113, 114] and by analysing the melting temperatures of short duplexes adjacent to hairpins [115].

The coarse-grained DNA model was derived in a "top-down" fashion, i.e. by choosing a physically motivated functional form, and then focusing on correctly reproducing the free energy differences between different states of the system, as opposed to a "bottom-up" approach that starts from a more detailed representation of DNA and typically focuses on accurate representation of local structural details. The interactions were originally fitted to reproduce melting temperatures of 'average' oligonucleotides, obtained by averaging over the parameters of SantaLucia's model [116]. A new, sequence-dependent version has been recently developed [117]. In addition, the model is fitted to reproduce the structural and mechanical properties of double- and single-stranded DNA such as the persistence length and the twist-modulus. The model allows for base pairing only between Watson-Crick complementary bases, but otherwise does not distinguish between bases in terms of interaction strengths.

The model was fitted to reproduce DNA behaviour at a salt concentration ($[Na^+] = 0.5\,M$) where the electrostatic properties are strongly screened, and it may be reasonable to incorporate them into a short-ranged excluded volume. Such high salt concentrations are typically used in DNA nanotechnology applications, hence motivating this approach. It should be noted that the model neglects several features of the DNA structure and interactions due to the high level of coarse-graining. Specifically, the double helix in the model is symmetrical rather than the grooves between the backbone sites along the helix having different sizes, and all four nucleotides have the same structure.

To cope with the complexity of the model and the large number of nucleotides involved in bulk simulations, we employ the GPU code presented in Chapter 6. As shown in Fig. 6.5, harvesting the power of modern Graphical Processing Units (GPUs) results in a 30-fold speed-up. The CPU version of the code, as well as the Python library written to simplify generation of initial configurations and post-processing analysis, have been released as free software [117, 118].

1.2.2 Bulk simulations

To compare numerical results with theory, we perform Brownian dynamics simulations of 400 dsDNA molecules made up by 24 nucleotides each, i.e. 400 cylinder-like objects with an aspect ratio of ≈ 2 (see Fig. 1.1 (d)). The integration time step has been chosen to be 0.003 in simulation units which corresponds, if rescaled with the units of length, mass and energy used in the model, to approximately 1×10^{-14} seconds.

We study systems at three different temperatures, namely $T = 270\,\mathrm{K}$, $285\,\mathrm{K}$ and $300\,\mathrm{K}$, and for different concentrations, ranging from $2\,\mathrm{mg/ml}$ to $241\,\mathrm{mg/ml}$. The $T = 270\,\mathrm{K}$ state point, despite being far from the experimentally accessed T, is here investigated to test the theory in a region of the phase diagram where the degree of association is significant. To quantify the aggregation process we define two DNADs as bonded if their pair interaction energy is negative. Depending on temperature and concentration, we use $10^6 - 10^7$ MD steps for equilibration and $10^8 - 10^9$ MD steps for data generation on NVIDIA Tesla C2050 GPUs, equivalent to $1 - 10\ \mu s$.

1.3 Theory

We build on the theoretical framework previously developed to account for the linear aggregation and collective ordering of quasi-cylindrical particles [52]. Here, we provide a discussion of how such a theory can be used to describe the reversible chaining and ordering of oligomeric DNADs at the level of detail adopted by the present model. According to Ref. [52], the free energy of a system of equilibrium polymers can be written as

$$\begin{aligned} \frac{\beta F}{V} &= \sum_{l=1}^{\infty} \nu(l) \left\{ \ln \left[v_d \nu(l) \right] - 1 \right\} + \\ &+ \frac{\eta(\phi)}{2} \sum_{\substack{l=1 \\ l'=1}}^{\infty} \nu(l) \nu(l') v_{excl}(l, l') \\ &- \beta \Delta F_b \sum_{l=1}^{\infty} (l-1) \nu(l) + \sum_{l=1}^{\infty} \nu(l) \sigma_o(l) \end{aligned} \tag{1.2}$$

where V is the volume of the system, v_d is the volume of a monomer, $\phi \equiv v_d \rho$ ($\rho = N/V$ is the number density of monomers) is the packing fraction, $\nu(l)$ is the number density of chains of length l, normalized such that $\sum_{l=1}^{\infty} l\, \nu(l) = \rho$, ΔF_b, as discussed in Section 1.4.2, is a parameter which depends on the free

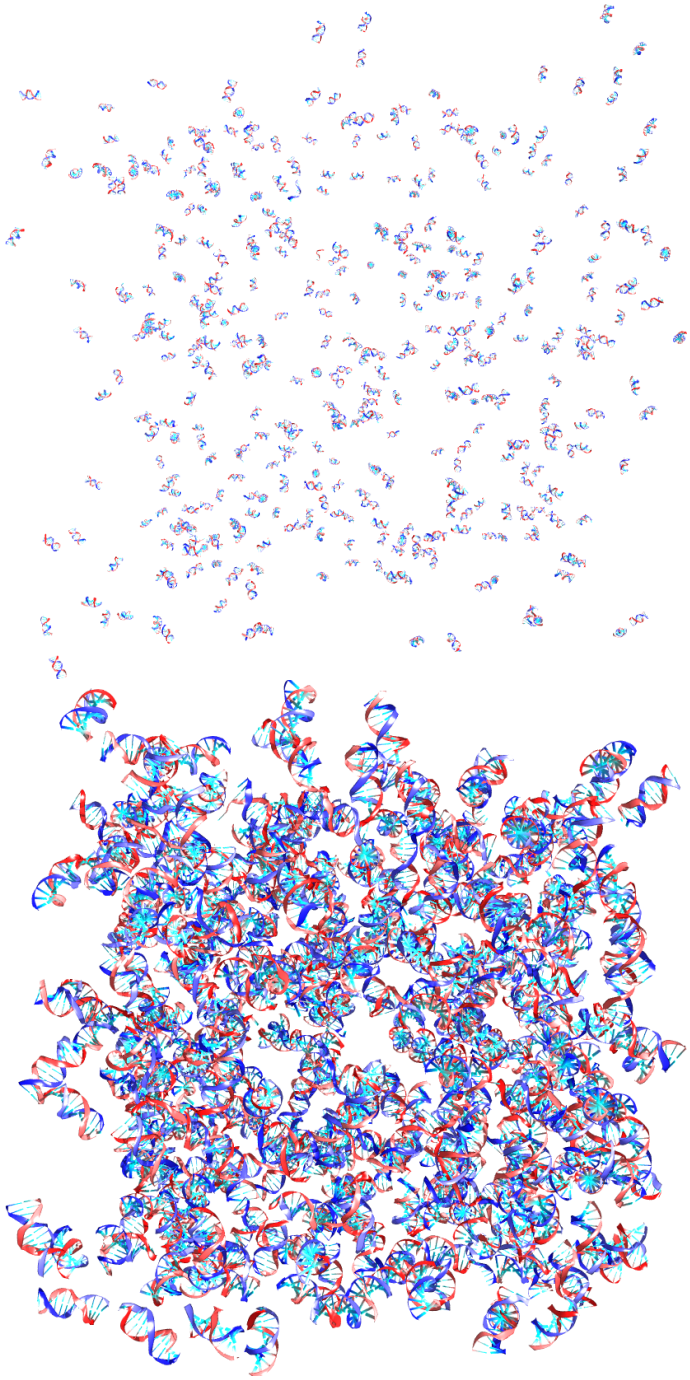

Figure 1.2: Snapshots taken from simulations at $T = 300\,\mathrm{K}$. At low concentrations ($c = 2\,\mathrm{mg/ml}$, top) chain formation is negligible and the average chain length is approximately 1. As the concentration is increased ($c = 80\,\mathrm{mg/ml}$, bottom), DNADs start to self-assemble into chains and the average chain length increases.

energy associated to a single bond and $v_{excl}(l, l')$ is the excluded volume of two chains of length l and l'. $\eta(\phi)$ is the Parsons-Lee factor [119]

$$\eta(\phi) = \frac{1}{4}\frac{4-3\phi}{(1-\phi)^2} \tag{1.3}$$

and $\sigma_o(l)$ [103] accounts for the orientational entropy that a chain of length l loses in the nematic phase (including possible contributions due to its flexibility). The explicit form for $\sigma_o(l)$ can be found in Ref. [52].

The free energy functional presented in Eq. (1.2) explicitly accounts for the polydispersity inherent in the equilibrium polymerization process using a discrete chain length distribution and for the entropic and energetic contributions of each single bond through the parameter ΔF_b.

1.3.1 Isotropic phase

In the isotropic phase, $\sigma_0 = 0$ and the excluded volume can be written as

$$v_{excl}(l, l', X_0) = 2B_I X_0^2 \, l \, l' + 2v_d k_I \frac{l + l'}{2} \tag{1.4}$$

where the parameters B_I and k_I can be estimated via MC integrals of a system composed by only two monomers (see Section 1.4.2) and X_0 is the aspect ratio of the monomers. We assume that the chain length distribution $\nu(l)$ is exponential [52] with an average chain length M

$$\nu(l) = \rho M^{-(l+1)}(M-1)^{l-1} \tag{1.5}$$

where

$$M = \frac{\sum_{l=1}^{\infty} l\,\nu(l)}{\sum_{l=1}^{\infty} \nu(l)}. \tag{1.6}$$

With this choice for $\nu(l)$ the free energy in Eq. (1.2) becomes:

$$\begin{aligned} \frac{\beta F_I}{V} &= -\rho\beta\Delta F_b(1 - M^{-1}) + \\ &+ \eta(\phi)\left[B_I X_0^2 + \frac{v_d k_I}{M}\right]\rho^2 + \\ &+ \frac{\rho}{M}\left[\ln\left(\frac{v_d\rho}{M}\right) - 1\right] + \\ &+ \rho\frac{M-1}{M}\ln(M-1) - \rho\ln M. \end{aligned} \tag{1.7}$$

Minimization of the free energy in Eq. (1.7) with respect to M provides the

following expression for the average chain length M:

$$M = \frac{1}{2}\left(1 + \sqrt{1 + 4\phi e^{k_1 \phi \eta(\phi) + \beta \Delta F_b}}\right). \tag{1.8}$$

1.3.2 Nematic phase

In the nematic phase the monomer orientational distribution function $f(\theta)$ depends explicitly on the angle θ between the particle and the nematic axis, i.e. the direction of average orientation of the DNAD, since the system is supposed to have azimuthal symmetry around such axis. We assume the form proposed by Onsager [45], i.e.:

$$f_\alpha(\theta) = \frac{\alpha}{4\pi \sinh \alpha} \cosh(\alpha \cos\theta) \tag{1.9}$$

where α controls the width of the angular distribution. Its equilibrium value is obtained by minimizing the free energy with respect to α. As discussed in Section 1.4.2, we assume the following form for the excluded volume in the nematic phase:

$$v_{excl}(l, l', X_0, \alpha) = 2B_N(\alpha) X_0^2 l\, l' + 2v_d k_N^{HC}(\alpha) \frac{l + l'}{2} \tag{1.10}$$

where the term $2v_d k_N^{HC}(\alpha)$ is the end-midsection contribution to the excluded volume of two hard cylinders (see Section 1.4.3) and

$$B_N(\alpha) = \frac{\pi}{4} D^3 \left(\eta_1 + \frac{\eta_2}{\alpha^{1/2}} + \frac{\eta_3}{\alpha}\right) \tag{1.11}$$

where D is the diameter of the monomer and η_k with $k = 1, 2, 3$ are three parameters that we chose in order to reproduce the excluded volume calculated from MC calculations as discussed in Section 1.4.2.

Inserting Eqs. (1.10) and (1.5) into Eq. (1.2), and assuming once more an exponential distribution for $\nu(l)$, one obtains after some algebra:

$$\begin{aligned}\frac{\beta F_N}{V} &= \hat{\sigma}_o - \rho\beta\Delta F_b(1-M^{-1}) + \\ &+ \eta(\phi)\left[B_N(\alpha)X_0^2 + \frac{v_d k_N^{HC}(\alpha)}{M}\right]\rho^2 + \\ &+ \frac{\rho}{M}\left(\ln\left[\frac{v_d\rho}{M}\right]-1\right) - \rho\ln M + \\ &+ \rho\ln(M-1)\frac{M-1}{M} \end{aligned} \tag{1.12}$$

where $\hat{\sigma}_o \equiv \sum_l \sigma_o(l)\nu(l)$. The explicit calculation of the parameters B_N and k_N^{HC} is explained in Sections 1.4.2 and 1.4.3.

Assuming that the orientational entropy $\hat{\sigma}_o$ can be approximated with the expression valid for long chains [103], minimization with respect to M results in

$$M = \frac{1}{2}\left(1+\sqrt{1+\alpha\phi e^{k_N(\alpha)\phi\eta(\phi)+\beta\Delta F_b}}\right). \tag{1.13}$$

while using the approximated expression for short chains [103], one obtains

$$M = \frac{1}{2}\left(1+\sqrt{1+4\alpha\phi e^{k_N(\alpha)\phi\eta(\phi)+\beta\Delta F_b-1}}\right). \tag{1.14}$$

The equilibrium value of α is thus determined by further minimizing the nematic free energy in Eq. (1.12), which has become only a function of α. The parameter α is related to the degree of orientational ordering in the nematic phase as expressed by the nematic order parameter S as follows:

$$S(\alpha) = \int (3\,\cos^2\theta - 1) f_\alpha(\theta)\pi\,\sin\theta\,d\theta \approx 1 - 3/\alpha. \tag{1.15}$$

Further refinements of the theory may be obtained by including a more accurate description of the orientational distribution $f_\alpha(\theta)$ in the proximity of the I-N phase transition, along the lines of Eqs. (40)-(42) of Ref. [52]. For the sake of simplicity we have just presented the basic theoretical treatment. However, in the theoretical calculations in Section 1.4 we will make use of the refined and more accurate free energy functional proposed in Ref. [52]. This upgraded version makes use of the fact that, in the nematic phases, the angular distribution of the monomers as computed in simulations of super-quadrics is isotropic and not Onsager-like, i.e. Eq. (1.9) becomes l-dependent. Although this change does not modify the qualitative behaviour of the calculated quantities, it improves the agreement with the numerical results.

1.3.3 Phase coexistence

The phase boundaries, at which the aggregates of DNAD are sufficiently long to induce macroscopic orientational ordering, are characterized by coexisting isotropic and nematic phases in which the volume fraction of DNADs are, respectively, $\phi_N = v_d \rho_N$ and $\phi_I = v_d \rho_I$. The number densities ρ_I and ρ_N can be calculated by minimizing Eq. (1.7) with respect to M_I and by minimizing Eq. (1.12) with respect to M_N and α. In addition, the two phases must be at equal pressure, i.e. $P_I = P_N$, and chemical potential, i.e. $\mu_I = \mu_N$. These conditions yield the following set of equations:

$$\begin{aligned} \frac{\partial}{\partial M_I} F_I(\rho_I, M_I) &= 0 \\ \frac{\partial}{\partial M_N} F_N(\rho_N, M_N, \alpha) &= 0 \\ \frac{\partial}{\partial \alpha} F_N(\rho_N, M_N, \alpha) &= 0 \\ P_I(\rho_I, M_I) &= P_N(\rho_N, M_N, \alpha) \\ \mu_I(\rho_I, M_I) &= \mu_N(\rho_N, M_N, \alpha) \end{aligned} \tag{1.16}$$

1.4 Results

1.4.1 Properties of the model

To characterize structural and geometrical properties of monomers and aggregates, we analyze conformations of duplexes extracted from large-scale GPU simulations (see Fig. 1.2 for some snapshots).

In the following, the volume v_d occupied by a single DNAD of length $X_0 D$ and double helix diameter D ($D \simeq 2\,\mathrm{nm}$) will be considered as the volume of a cylinder with the same length and diameter, i.e. $v_d = \pi X_0 D^3/4$. When comparing numerical and experimental results with theoretical predictions we use the number of base pairs N_b in place of X_0 ($X_0 \simeq 0.172 N_b$) and the concentration c instead of the packing fraction ϕ, which can be related to the former via:

$$\phi = \frac{0.172 D^3 \pi}{8 m_N} c \tag{1.17}$$

where $m_N = 330\,\mathrm{Da}$ is the average mass of a nucleotide. Hence, in the following c_I and c_N will be used in place of ϕ_I and ϕ_N.

First we calculate the dimensions (height L and width D) of the DNADs for

different c and T. We observe no concentration dependence on both quantities, while the variation in T is negligible (of the order of 0.1% between DNADs of samples at 270 K and 300 K). The effect of this small change does not affect substantially the value of the aspect ratio, which we consider constant ($X_0 = 2.06$) throughout this work.

The geometrical properties of end-to-end bonded duplexes are not well-known since there are no experimental ways to probe such structures. In a very recent work, the interaction between duplex terminal base–pairs has been analyzed by means of large-scale full-atom simulations by Maffeo *et al.* [120]. They found that blunt-ended duplexes (i.e. duplexes without dangling ends) have preferential binding conformations with different values of the azimuthal angle γ, defined as the angle between the projections onto the plane orthogonal to the axis of the double helix of the vectors connecting the O5′ and O3′ terminal base pairs. They report two preferential values for γ, namely $\gamma = -20^\circ$ and $\gamma = 180^\circ$.

In the present model the continuity of the helix under end-to-end interactions is intrinsic in the model and the azimuthal angle probability distribution is peaked around a single value $\gamma_0 \approx 40^\circ$ (see Fig. 1.3(a)). This is very close to the theoretical value $\gamma \approx 36^\circ$ given by the pitch of the B-DNA double helix. The qualitative difference between conformations of bonded DNADs found in this work and in Ref. [120] should be addressed in future studies, in order to describe the coaxial end-to-end interaction in a more proper way.

In addition, we calculate the average distance r between the centres of mass of the terminal base pairs. Fig. 1.3(b) shows $P(r)$, the probability distribution of r. $P(r)$ is peaked at 0.39 nm, whereas Maffeo *et al.* [120] found an average distance of $r \approx 0.5$ nm. This difference can be understood in terms of the effect of the salt concentration which, being five times higher than the one used in Ref. [120], increases the electrostatic screening, thus effectively lowering the repulsion between DNA strands. The sharp minimum which occurs at $r \approx 0.39$ for every temperature seems to be a feature of the system and hence its origin could be traced back to the shape of the interaction term acting between the two ending base-pairs. However, we can not rule out the possibility that it may be due to numerical noise.

The effect of the temperature is small, as lowering T leads only to more peaked distributions for both $P(\gamma)$ and $P(r)$ (and a very small shift towards smaller angles for γ) but does not change the overall behaviour.

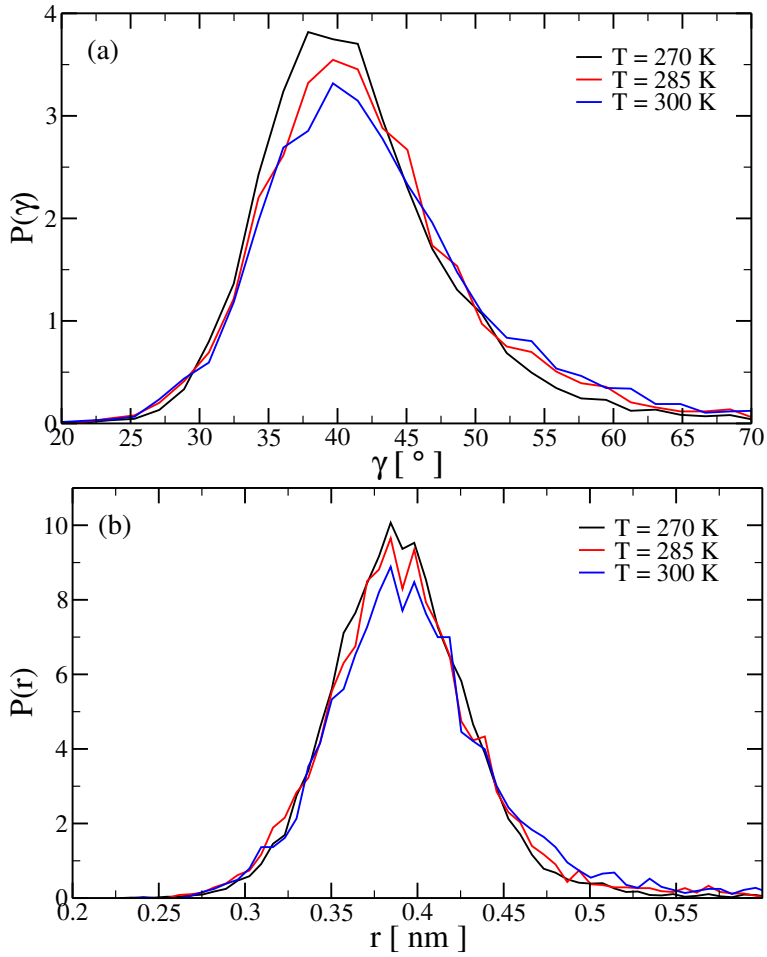

Figure 1.3: Probability distributions for (a) the azimuthal angle γ and (b) the end-to-end distance r.

1.4.2 Stacking free energy and excluded volume

In this section we discuss the procedure employed to evaluate the input quantities required by the theory, namely ΔF_b and $v_{excl}(l, l')$. To this aim we perform a Monte Carlo integration over the degrees of freedom of two duplexes. ΔF_b is defined as [52]

$$\beta \Delta F_b = \ln \left[2 \frac{\Delta(T)}{v_d} \right] \tag{1.18}$$

where [37]

$$\Delta(T) = \frac{1}{4} \left\langle \int_{V_b} [e^{-\beta V(\mathbf{r}_{12}, \mathbf{\Omega}_1, \mathbf{\Omega}_2)} - 1] \, d\mathbf{r}_{12} \right\rangle . \tag{1.19}$$

Here $\mathbf{r}_{12}$ is the vector joining the centres of mass of particles 1 and 2, $\mathbf{\Omega}_i$ is the orientation of particle i and $\langle \ldots \rangle$ represents an average taken over all possible

orientations. V_b is the bonding volume, defined here as the set of points where the interaction energy $V(\mathbf{r}_{12}, \mathbf{\Omega}_1, \mathbf{\Omega}_2)$ between duplex 1 and duplex 2 is less than $k_B T$. To numerically evaluate $\Delta(T)$ we perform a MC integration using the following scheme:

1. Produce an ensemble of 500 equilibrium configurations of a single duplex at temperature T.

2. Set the counter $N_{\text{tries}} = 0$ and the energy factor $F = 0$.

3. Choose randomly two configurations i and j from the generated ensemble.

4. Insert a randomly oriented duplex i in a random position in a cubic box of volume $V = 1000\ nm^3$. Insert a second duplex j in a random position and with a random orientation. Compute the interaction energy $V(i,j)$ between the two duplexes i and j and, if $V(i,j) < k_B T$, update the energy factor, $F = F + \left(e^{-\beta V(i,j)} - 1\right)$. Increment N_{tries}.

5. Repeat from step 3, until $\Delta(T) \cong \frac{1}{4}\frac{V}{N_{\text{tries}}}F$ converges within a few per cent precision.

The employed procedure to compute $v_{excl}(l, l')$ is fairly similar except that it is performed for duplexes with a various number of bases (i.e. with different X_0) and the quantity F counts how many trials originate a pair configuration with $V(i,j) > k_B T$ (i.e. in step 4, $F = F + 1$). In the nematic case, the orientations of the duplexes are extracted randomly from the Onsager distribution given by Eq. (1.9). With such procedure,

$$v_{excl}(l = 1, l' = 1, X_0) = \frac{V}{N_{\text{tries}}}F \tag{1.20}$$

We calculate v_{excl} for 8 values of α, ranging from 5 to 45 (see Section 1.4.3). Since the X_0 and l dependences of Eqs. (1.4) and (1.10) are the same and the X_0 dependence of the numerically calculated v_{excl} on the shape of DNADs is negligible, the evaluation of the excluded volume as a function of X_0 provides the same information as the evaluation of v_{excl} as a function of l.

We have checked that the dependence of $\Delta(T)$ and $v_{excl}(l, l')$ on the energy threshold employed in step 4 is negligible.

Fig. 1.4 shows $\Delta(T)$ for all investigated T in a $\ln\Delta$ vs $1/T$ plot. A linear dependence properly describes the data at the three T. An alternative way to evaluate $\Delta(T)$ is provided by the limit $\rho \to 0$ of Eq. (1.8). Indeed in the low

density limit M and $\Delta(T)$ are related via the following relation:

$$\Delta(T) = \frac{M(1-M)}{2\rho}. \tag{1.21}$$

Therefore it is also possible to estimate $\Delta(T)$ by extrapolating the low density data for M at $T = 270\,\mathrm{K}$, $285\,\mathrm{K}$ and $300\,\mathrm{K}$. The results, also shown in Fig. 1.4, are in line with the ones obtained through MC calculations. The Arrhenius behavior of $\Delta(T)$ suggests that bonding entropy and stacking energy are in first approximation T independent. The coaxial stacking free energy G_{ST} is related to $\Delta(T)$ as follows

$$G_{ST} = -k_B T \ln[2\rho\Delta(T)]. \tag{1.22}$$

Substituting the fit expression provided in Fig. 1.4 for $\Delta(T)$ results in a stacking free energy $G^0_{ST} = -0.086\,\mathrm{kcal/mol}$ at a standard concentration 1 M of DNADs and $T = 293\,\mathrm{K}$ comprising a bonding entropy of $-30.6\,\mathrm{cal/mol\,K}$ and a bonding energy of $-9.06\,\mathrm{kcal/mol}$.

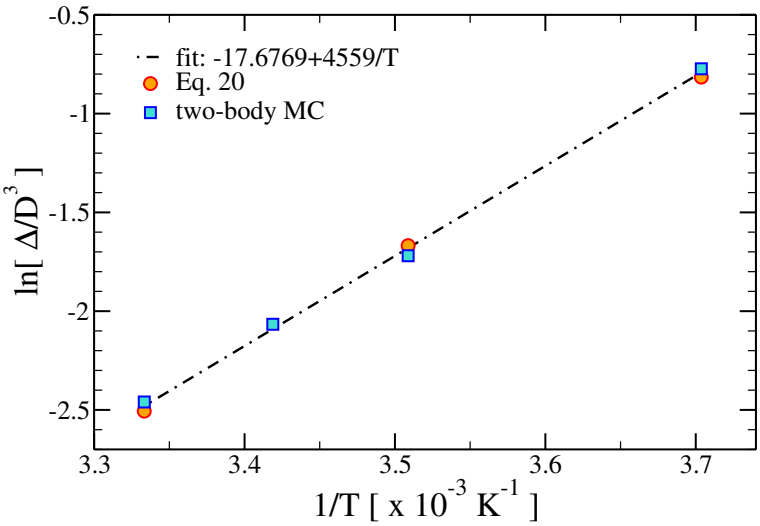

Figure 1.4: $\Delta(T)$ calculated with the procedures described in Section 1.4.2 for all investigated T.

Following Ref. [52], the excluded volume is assumed to be the following

second order polynomial in l and l':

$$v_{excl}[l,l';f(\theta)] = 2\int f(\theta)f(\theta')D^3\left[\Psi_1(\gamma,X_0)+ \right.$$
$$\left.+\frac{l+l'}{2}\Psi_2(\gamma,X_0)X_0+\Psi_3(\gamma,X_0)X_0^2\,l\,l'\right]$$
$$d\mathbf{\Omega}\,d\mathbf{\Omega}' \tag{1.23}$$

where the functions Ψ_α, $\alpha = 1,2,3$, describe the angular dependence of the excluded volume. The orientational probability $f(\theta)$ is normalized such that

$$\int f(\theta)d\mathbf{\Omega} = 1. \tag{1.24}$$

The three contributions to the excluded volume in Eq. (1.23) come from end-end, end-midsection and midsection-midsection steric interactions [52] between two chains.

In the isotropic phase the orientational distribution does not have any angular dependence, i.e. $f(\theta) = 1/4\pi$, and Eq. (1.23) reduces to the form

$$v_{excl}(l,l',X_0) \quad = \quad B_I X_0^2\,l\,l' + k_I v_d\frac{l+l'}{2} + A_I. \tag{1.25}$$

The parameters B_I, k_I and A_I appearing in Eq. (1.25) can be calculated via MC integration procedures as discussed previously and in Ref. [52]. We expect that these parameters do not depend on X_0 because each DNAD comprises N_b stacked base pairs which are all identical with respect to excluded volume interactions (i.e. they all have the same shape). In particular, the calculated excluded volume of two DNADs is reported in Fig. 1.5 for 5 different aspect ratios, together with the resulting values for the above parameters.

Using the Onsager angular distribution $f_\alpha(\theta)$ in Eq. (1.9), the excluded volume in the nematic phase depends also on the parameter α, i.e. the general form in Eq. (1.23) reduces to

$$\begin{aligned} v_{excl}(l,l',X_0,\alpha) \quad &= \quad B_N(\alpha)X_0^2\,l\,l' + k_N(\alpha)v_d\frac{l+l'}{2} \\ &+ \quad A_N(\alpha). \end{aligned} \tag{1.26}$$

Assuming that $A_N(\alpha) = 0$, $k_N(\alpha) = k_N^{HC}(\alpha)$ and $B_N(\alpha)$ is given by Eq. (1.11), the three parameters η_k with $k = 1,2,3$ have to be estimated. For $l = l' = 1$ and several values of α ($\alpha = 5\ldots45$ in steps of 5) and X_0 we calculated numerically the nematic excluded volume for two DNADs. The results are shown in Fig. 1.6,

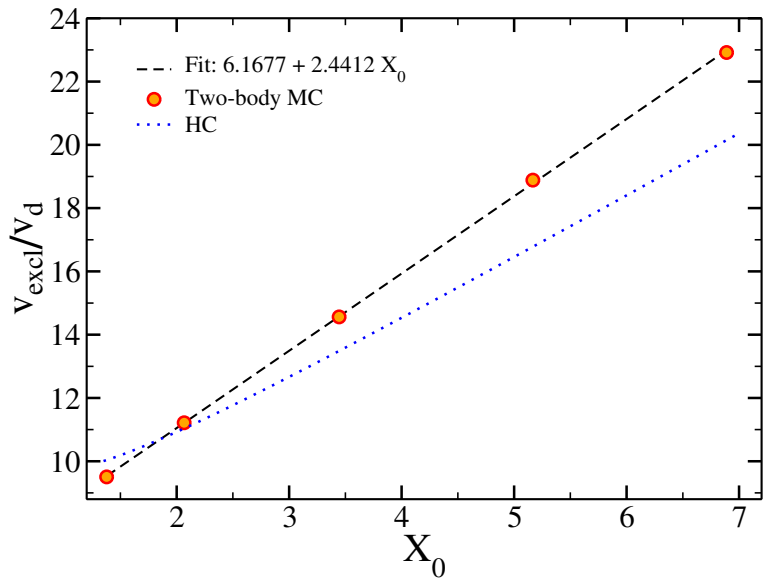

Figure 1.5: Excluded volume in the isotropic phase together with analytic approximations. From the linear fit one has $B_I = 0.959D^3$ and $k_I = 3.084$, while we assume $A_I = 0$

where we plot v_{excl}/v_d vs X_0 for various α. The dashed lines shown in Fig. 1.6 are obtained through a two-dimensional fit to numerical data for $v_{excl}(1, 1, X_0, \alpha)$ using Eq. (1.10) as fitting function.

1.4.3 Excluded volume of hard cylinders

For two rigid chains of length l and l' which are composed of hard cylinders (HCs) of diameter D and length $X_0 D$, $v_{excl}(l, l')$ can be described by

$$\begin{aligned} v_{excl}^{HC}[l, l'; f(\theta)] &= \int f(\theta) f(\theta') D^3 \Big[\frac{\pi}{2} \sin\gamma + \frac{\pi}{2} X_0 \\ &\quad (1 + |\cos\gamma| + \frac{4}{\pi} E(\sin\gamma)) + \\ &+ \quad \frac{l + l'}{2} + 2X_0^2 \sin\gamma \;\; l\, l' \Big] \, d\mathbf{\Omega}\, d\mathbf{\Omega}' \end{aligned} \tag{1.27}$$

where $\cos\gamma = \mathbf{u} \cdot \mathbf{u}'$, $\mathbf{u}$ and $\mathbf{u}'$ are the orientations of two HCs and $E(\sin\gamma)$ is the complete elliptical integral

$$E(\sin\gamma) = \frac{1}{4} \int_0^{2\pi} (1 - \sin^2\gamma \sin^2\psi)^{1/2} d\psi. \tag{1.28}$$

The integrals in Eq. (1.27) can be calculated exactly in the isotropic phase,

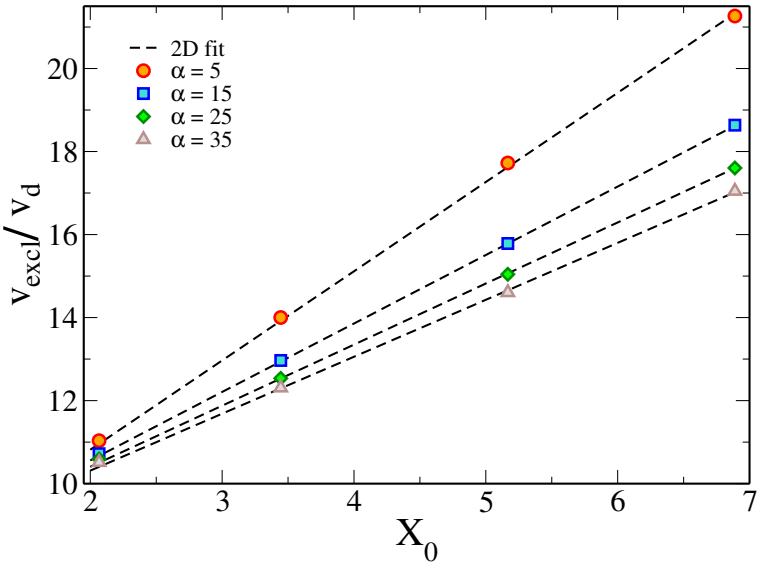

Figure 1.6: Excluded volume as a function of aspect ratio X_0 in the nematic phase together with analytic approximations for several α. The dashed lines are obtained plotting the function reported in Eq. (1.10) and setting $\eta_1 = 0.386419$, $\eta_2 = 1.91328$ and $\eta_3 = -0.836354$.

while in the nematic phase the calculation can be done analytically only for suitable choices of the angular distribution $f(\theta)$. Here we assume that the angular distribution is given by the Onsager function in Eq. (1.9).

Using the Onsager orientational function the following approximate expressions for the coefficients $k_N(\alpha)$, $B_N(\alpha)$ and $A_N(\alpha)$ can be derived [103]

$$\begin{aligned} \tilde{B}_N(\alpha) &= D^3(\pi/4)\rho_a(\alpha) \\ \tilde{k}_N(\alpha) &= \pi D^3 \frac{X_0}{v_d}\left(1-\frac{1}{\alpha}\right) \\ \tilde{A}_N(\alpha) &= D^3\left(\pi/4\right)^2\rho_a(\alpha) \end{aligned} \tag{1.29}$$

where

$$\rho_a = 4(\pi\alpha)^{-1/2}\left(1-\frac{15}{16\,\alpha}+\frac{105}{512\,\alpha^2}+\frac{315}{8192\,\alpha^3}\right). \tag{1.30}$$

We evaluate numerically the excluded volume in Eq. (1.27) for many values of α and, building on the expressions in Eqs. (1.29), we perform a fit to this data

using the following functions:

$$B_N^{HC}(\alpha) \simeq D^3(\pi/4)\left(\rho_a(\alpha) + \frac{c_4}{\alpha^{9/2}} + \frac{c_5}{\alpha^{11/2}}\right) \tag{1.31}$$

$$k_N^{HC}(\alpha) = 4\left(1 - \frac{1}{\alpha}\right) + \sum_{i=2}^{\infty} \frac{b_i}{\alpha^i} \simeq \frac{4}{\pi} \sum_{i=0}^{4} \frac{d_i}{\alpha^i} \tag{1.32}$$

$$A_N^{HC}(\alpha) \simeq D^3(\pi/4)^2\left(\rho_a(\alpha) + \frac{c_4}{\alpha^{9/2}} + \frac{c_5}{\alpha^{11/2}}\right) \tag{1.33}$$

The coefficient values resulting from the fitting procedure are $c_4 = 1.2563$, $c_5 = -0.95535$, $d_0 = 3.0846$, $d_1 = -4.0872$, $d_2 = 9.0137$, $d_3 = -9.009$ and $d_4 = 3.3461$.

1.4.4 Isotropic phase: comparing simulation results with theoretical predictions

Fig. 1.7 shows the concentration c dependence of M calculated from the MD simulation of the $N_b = 12$ system. The average chain length increases progressively on increasing c. The figure also shows the theoretical predictions calculated by minimizing the isotropic free energy in Eq. (1.7) with respect to M using the previously discussed estimates for ΔF_b and v_{excl}. The theoretical results properly describe the MD simulation data up to concentrations around 200 mg/ml, which corresponds to a volume fraction $\phi \approx 0.20$. In Ref. [52] similar observations have been made and the discrepancy at moderate and high ϕ has been attributed to the inaccuracy of the Parsons decoupling approximation. The M values calculated using the excluded volume of two hard cylinders (HC) are also reported, to quantify the relevance of the actual shape of the DNA duplex. Indeed the HC predictions appreciably deviate from numerical data beyond 100 mg/ml.

1.4.5 Phase coexistence: theoretical predictions

A numerical evaluation of the phase coexistence between the isotropic and the nematic phases for the coarse-grained model adopted in this study is still impossible to obtain given the current computational power. We thus limit ourselves to the evaluation of the I-N phase coexistence via the theoretical approach discussed in Section 1.3. Fig. 1.8 shows the theoretical phase diagram in the c-N_b plane for $T = 270$ K and 300 K. As expected, both c_I and c_N decrease on increasing N_b, since the increase of the number of bases results in a larger aspect ratio which, as showed by Onsager, controls the difference between the

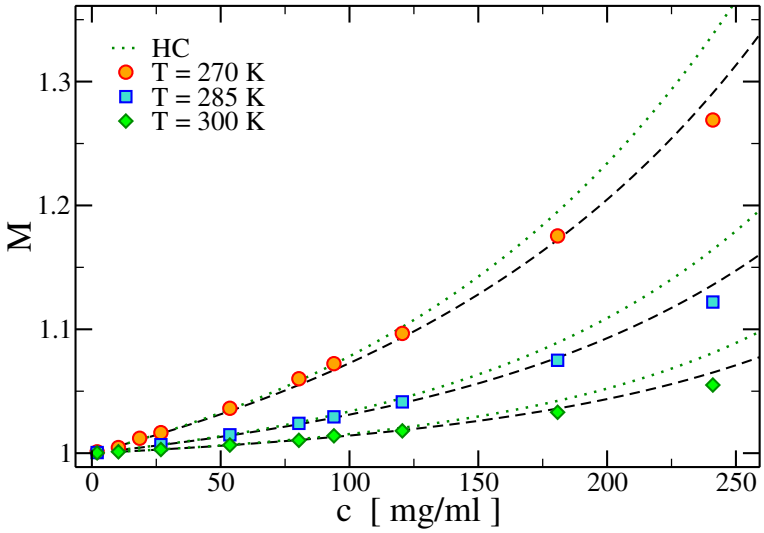

Figure 1.7: Average chain length M in the isotropic phase at low concentration. Symbols are numerical results and dashed lines are theoretical predictions. Dotted lines are theoretical predictions using the excluded volume of HCs v_{excl}^{HC} (see Section 1.4.3).

excluded-volume terms in the isotropic and nematic free energies. However, as it will be discusses in Section 1.4.7, mapping a self-assembling system onto a mono-disperse system by using the average chain length as an effective aspect ratio can yield qualitatively incorrect results. On decreasing T, theory predicts a 10% decrease of c_I and a similar decrease of c_N, resulting in an overall shift of the I-N coexistence region toward lower c values. This trend is related to the increase of the average chain length M with increasing $\beta\Delta F_b$ (see Fig. 1.4).

Fig. 1.8 also shows the phase boundaries calculated using the excluded volume of two hard cylinders. Assimilating DNADs to hard cylinders results in a 10–15% widening of the isotropic-nematic coexistence region.

1.4.6 Comparison between theory and experiments

The theoretical predictions concerning the isotropic-nematic coexisting concentrations can be compared to the experimental results reported in Refs. [49] and [91] for blunt-ended DNADs.

Fig. 1.9 compares the experimentally determined nematic concentrations c_N at coexistence with the values calculated from the present model for $T = 293\,\mathrm{K}$. Despite all the simplifying assumptions and despite the experimental uncertainty, the results provide a reasonable description of the N_b dependence

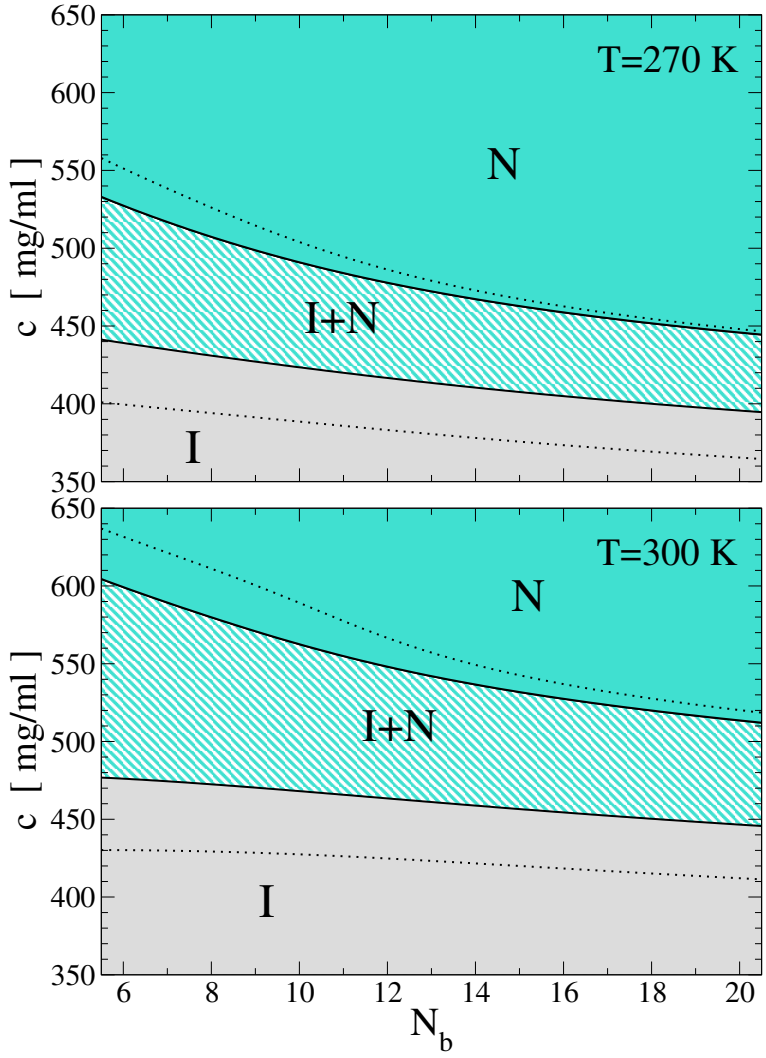

Figure 1.8: I-N phase diagram in the c vs N_b plane for $T = 270\,\mathrm{K}$ (top) and $300\,\mathrm{K}$ (bottom). Dotted lines are theoretical phase boundaries calculated using the excluded volume of HCs v_{excl}^{HC} (see Section 1.4.3).

of c_N. The experimental data refer to different base sequences and different salt concentrations. According to the authors, c_N is affected by an error of about $\pm 50\,\mathrm{mg/ml}$. In particular, for the case $N_b = 12$ the critical concentrations c_N for distinct sequences show that blunt-end duplexes of equal length but different sequences can display significantly different transition concentrations. Hence, for each duplex length, we consider the lowest transition concentration among the ones experimentally determined, since this corresponds to the sequence closest to the symmetric monomer in the model. Indeed the dependence of c_N on the DNADs sequence is expected to be larger for the shortest sequences, i.e.

$N_b < 12$, for which DNAD bending could be significant [91]. Unfortunately, quantitative experimental data on this bending effect are still lacking. In general it is possible that c_N for $N_b < 12$ (for which a large number of sequences have been studied, see Fig. 1.9), would be corrected to lower values if a larger number of sequences were explored. For more details on this phenomenon, we refer the reader to the discussions in Refs. [52, 91, 121].

The overestimation of the phase boundaries for $N_b \geq 12$ with respect to experimental results suggests that the DNA model of Ouldridge *et al.* [110] overestimates the coaxial stacking free energy. Such discrepancy can perhaps be attributed to the restricted number of microstates allowing for bonding states in the DNA model [110, 112], as discussed in Section 1.4. Indeed, allowing DNADs to form end-to-end bonds with more than one preferred azimuthal angle would increase the entropy of bonding, thus effectively lowering G_{ST}. Allowing for both left- and right-handed binding conformations, a possibility supported by the results of Maffeo *et al.* [120], would double $\Delta(T)$ in Eq. (1.22) and hence add an entropic contribution equal to $-k_B T \ln(2)$ to G_{ST}, which would result in a value $G^0_{ST} - 0.403\,\text{kcal/mol} = -0.49\,\text{kcal/mol}$ for $T = 293K$. Fig. 1.9 also shows the theoretical predictions for such upgraded G_{ST} value.

In Fig. 1.9 theoretical calculations of the I-N transition lines are shown for $G_{ST} = -0.4\,\text{kcal/mol}$ and $G_{ST} = -2.4\,\text{kcal/mol}$ at $T = 293\,\text{K}$ as the upper and lower boundaries of the grey band respectively. To calculate these critical lines we retain the excluded volume calculated in Section 1.4.2 and, given the value of G_{ST}, we evaluate ΔF_b according to Eqs. (1.18) and (1.22) for $T = 293\,\text{K}$ and ρ corresponding to the standard 1 M concentration.

The selected points with $N_b \geq 12$ fall within the grey band shown in Fig. 1.9, enabling us to provide an indirect estimate of G_{ST} between $-0.4\,\text{kcal/mol}$ and $-2.4\,\text{kcal/mol}$. For the points with $N < 12$, where duplex bending might play a role, it would be valuable to have more experimental points corresponding to more straight sequences in order to validate the theoretical predictions.

It is worth observing that for all DNAD lengths N_b, the electrostatic interactions are properly screened. For $N_b = 20$ a concentration 1.2 M of NaCl has been added to the solution resulting in a Debye screening length $k_D^{-1} \approx 0.23\,\text{nm}$. For all other lengths (i.e. $N_b \leq 18$) we note that at the lowest DNA concentration of 440 mg/ml corresponding to $N_b = 14$, $k_D^{-1} \approx 0.40\,\text{nm}$. Therefore the experimental k_D^{-1} is always smaller than the excluded volume diameter for the backbone-backbone interaction of our coarse-grained model [110] ($\approx 0.6\,\text{nm}$), thus enabling us to neglect electrostatic interactions.

On the other hand, if electrostatic interactions are not properly screened the effective aspect ratio for such DNAD sequences would be smaller than the ones used in our theoretical treatment and this would result in a underestimate of c_N. To account for this behavior one should at least have a reasonable estimate of the effective size of DNADs when electrostatic interactions are not fully screened. Moreover, the role of electrostatics can be subtle and not completely accounted for by simply introducing an effective size of DNADs. A possible route to include electrostatics in our treatment can be found in Ref. [82] and it will be addressed in future studies.

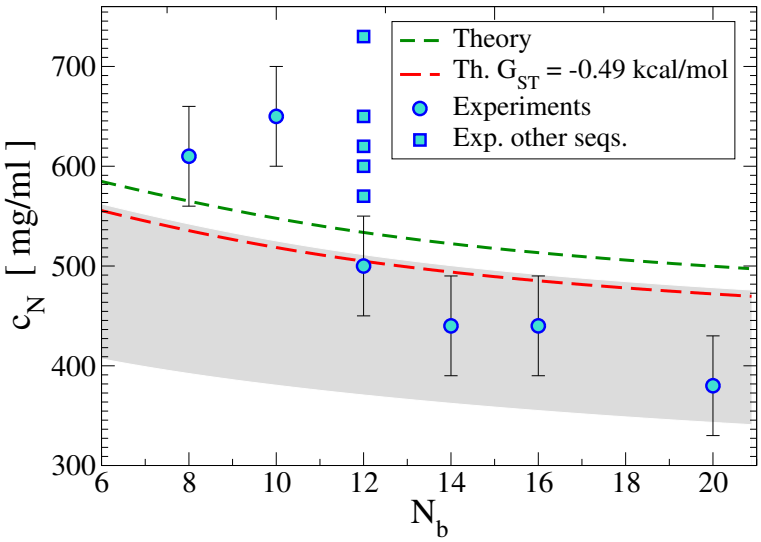

Figure 1.9: Critical nematic concentrations c_N as a function of the number of base pairs per duplex N_b for the present model, calculated theoretically at $T = 293\,\mathrm{K}$ using the computed stacking free energy G^0_{ST} (short dashed lines), $G_{ST} = -0.49\,\mathrm{kcal/mol}$ (long dashed lines), and for experiments [49] (circles and squares). Squares are c_N for different sequences at the same $N_b = 12$. The grey band has been built considering for G_{ST} an upper bound of $-0.4\,\mathrm{kcal/mol}$ and a lower bound of $-2.4\,\mathrm{kcal/mol}$.

1.4.7 Comparison with Onsager theory

The experimental average aggregation numbers are estimated in Refs. [49,75] by mapping the self-assembled system onto an "equivalent" mono-disperse system of hard rods with an aspect ratio equal to MX_0. In Ref. [52] it has been shown that the theoretically estimated isotropic-nematic coexistence lines for the case of polymerizing superquadric particles in the $MX_0 - \phi$ plane, parametrized by the stacking energy, are significantly different from the corresponding Onsager

original predictions (as re-evaluated in Ref. [95]). In light of the relevance for interpreting the experimental data, we show in Fig. 1.10 the same curves for the DNA model investigated here. In this model, a clear re-entrant behavior of the transition lines in the $c - MX_0$ plane is observed. The re-entrant behavior occurs for values of the stacking free energy accessed at temperatures between 270 K and 330 K and it arises as a result of the competition between steric (entropic) and temperature (energetic) contributions to the free energy in driving the self-assembly process which leads the system to the isotropic-nematic transition. We believe that the re-entrancy of the transition lines in the $c - MX_0$ plane is a peculiar mark of the system polydispersity, which results from the reversible self-assembly of chains.

1.5 Conclusions

In this article we have studied an equilibrium bulk solution of blunt ended DNA duplexes undergoing reversible self-assembly into chains, promoted by stacking interactions. The simulation study, carried out at different concentrations and temperatures, provides a clear characterization of the c and T dependence of the average polymerization length M and an indirect estimate of the stacking free energy. We have provided a theoretical description of the self-assembly process based on a theoretical framework recently developed in Ref. [52]. The inputs required by the theory (the DNAD excluded volumes and the stacking free energy) have been numerically calculated for the present DNA model, allowing for a parameter free comparison between molecular dynamic results and theoretical predictions. Such comparison has been limited to the isotropic phase, due to the difficulties in simulating the dense nematic phase under equilibrium conditions. The description of the isotropic phase is satisfactory: quantitative agreement between theory and simulations is achieved for concentrations up to $c \approx 200$ mg/ml. The stacking free energy value that properly accounts for the polymerization process observed in the molecular dynamics simulations is $G^0_{ST} = -0.086$ kcal/mol at a standard concentration 1 M of DNADs and $T = 293$ K comprising a bonding entropy of -30.6 cal/mol K and a bonding energy of -9.06 kcal/mol.

Theoretical predictions for the I-N transition have been compared with experimental results for several DNA lengths, ranging from 8 to 20 bases. For $N_b \geq 12$, the model predicts values for c_N which are higher than experimental ones. This suggests that the DNA model employed overestimates G_{ST}. In view

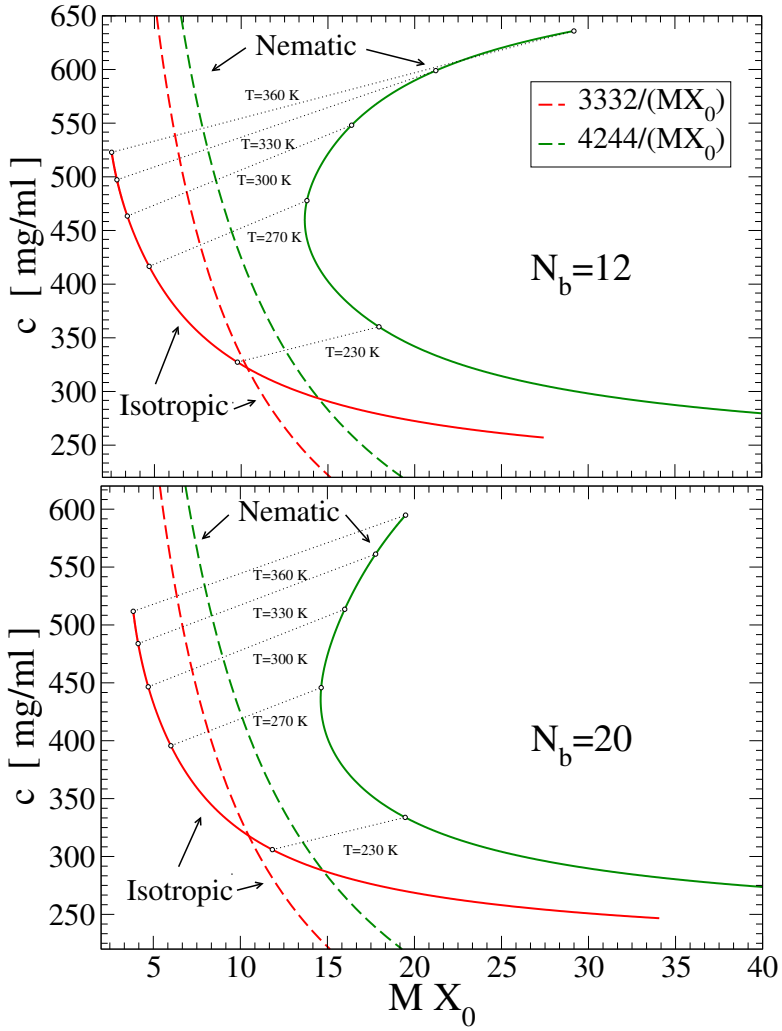

Figure 1.10: Isotropic-nematic coexistence lines in the average aspect ratio MX_0 and concentration c plane for two values of N_b, namely $N_b = 12$ (top) and $N_b = 20$ (bottom). Solid lines indicate theoretical predictions, dashed lines indicate Onsager original predictions, as re-evaluated in Ref. [95] for ϕ_I and ϕ_N and here reported in terms of the concentrations c_I and c_N expressed in mg/ml. Symbols along the isotropic and nematic phase boundaries at coexistence are joined by dotted lines, to indicate the change in concentration and average chain length at the transition.

of the recent results of Maffeo *et al.* [120], we speculate that the bonding entropy is underestimated, in agreement with the observation that the probability distribution of the azimuthal angle between two bonded DNADs, which is designed to be single-peaked, is too restraining. In this respect, the present study calls for an improvement of the coarse-grained potential [110] in regard to the coaxial stacking interaction.

The value of G_{ST} can also be used as a fitting parameter in the theory for matching c_N with the experimental results, retaining the excluded volume estimates calculated for the coarse-grained DNA model. Such procedure shows that values of the stacking free energy between $-0.4\,\mathrm{kcal/mol}$ and $-2.4\,\mathrm{kcal/mol}$ are compatible with the experimental location of the I-N transition line. In the work of Maffeo *et al.*, the authors report a more negative value of G_{ST}, namely $G_{ST}^{M} = -6.3\,\mathrm{kcal/mol}$, a value which was confirmed by the same authors by performing an investigation of the aggregation kinetic in a very lengthy all-atom simulation of DNAD with $N_b = 10$. If such G_{ST} value is selected as input in our theoretical approach (maintaining the same excluded volume term), then one finds $c_N^M \approx 250$ mg/ml, a value significantly smaller than the experimental result ($c_N = 650 \pm 50$ mg/ml). This casts some doubts on the effectiveness of the employed all-atom force-field to properly model coaxial stacking.

Finally, our work draws attention to the errors affecting the estimate of the average chain length M via a straightforward comparison of the nematic coexisting concentrations with analytic predictions based on the original Onsager theory for mono-disperse thin rods [49,75]. We have found that such approximation significantly underestimates M at the I-N transition concentration c_N. In addition, the theoretical approach predicts a re-entrant behavior of the transition lines in the c-MX_0 plane, a distinct feature of the polydisperse nature of the equilibrium chains.

Chapter 2

Mimicking low-density dipolar fluids: ring-forming bivalent patchy particles

2.1 Introduction

The equilibrium self-assembly of particles into chains and rings has been the subject of thorough investigations in the past [122–128]. In the limit of non-interacting aggregates (i.e. low densities) and weakly T-dependent persistence length, a well defined temperature exists below which a ring of a characteristic size becomes equally probable as the chain of the same size. For lower T, rings start to populate the system and, under appropriate conditions, may become the most abundant aggregates. The basic elements controlling the competition between rings and chains are well understood [129]. Compared to linear polymers, rings are energetically stabilized by the presence of an additional bond but pay a significant entropic cost, which can be split into two contributions, one proportional to the entropy loss associated to the necessity to constrain the end-to-end distance to distances of the order of the bonding length, the other proportional to the ring size arising from the number of ways a ring can open to form a chain. Despite the formal understanding, no quantitative parameter-free investigation of the ring-chain equilibrium for a well defined interaction potential has been reported so far. Two reasons suggest an effort in this direction:

- the Wertheim perturbation theory [130,131] has been shown to be a predictive and quantitative theory for describing chaining of large persistence polymers, when rings do not compete [132,133]. While attempts to generalize to the case in which rings are relevant have been presented [128,134–136], no quantitative comparison has been reported.

- The recent discovery that in the dipolar hard sphere (DHS) model a non negligible fraction of rings appears at low T, in the region in which a gas-liquid critical point was expected to occur [61,62] (see Chapter 3). Analogous findings have been documented in both experiments and simulations of colloids in which the dipolar interaction is dominant [137–141].

Wertheim theory, when applied to patchy colloidal particles with asymmetric interactions [56,57,142,143], provides a valuable model for interpreting the competition between chaining and branching, which is deemed to be essential in the physics of DHSs. It is thus important to extend it, under strict control, to the case in which rings are possible.

In this Chapter we present a two-patch model with controlled persistence length and introduce an extension of Wertheim theory to cope with the presence of rings. For one specific value of the opening bond angle (a parameter of the interaction potential) we then solve the Wertheim theory augmented with rings and compare theoretical predictions with simulation results in a large density region, from $\rho\sigma^3 = 5 \times 10^{-6}$ up to $\rho\sigma^3 = 0.02$, illuminating the presence of an optimal density for ring formation. The quality of the resulting parameter-free description provides a stringent test of Wertheim theory as a model of the thermodynamics of the self-assembly of equilibrium chains and rings. The augmented Wertheim theory, which is introduced in Section 2.3, has been developed by prof. J. M. Tavares.

2.2 Methods

We focus on a system of hard-sphere (HS) particles (of diameter σ, the unit of length) with two oppositely located patches. The patch-patch interaction is modelled via a Kern-Frenkel (KF) potential [144], i.e. a square-well interaction of range Δ and depth u_0, multiplied by an angular function which accounts for the orientational contribution, such that particles bind only when distinct patches face each other. In the standard KF model the single-bond per patch condition depends on the patch angular width $\theta_{\max}$ and Δ. The maximum

opening angle $\theta_{\max}$ must satisfy the condition $\sin(\theta_{\max}) \leq [2(1+\Delta/\sigma)]^{-1}$. In the limit of a cohesive contact potential ($\Delta \to 0$), the maximum opening angle is 30°, a value for which the probability of ring formation is minute. To be able to extend the range of values of $\theta_{\max}$ beyond 30° and to decouple the choice of the angle from the choice of the range we take inspiration from the N_{max} model [38, 145], imposing numerically in the Monte Carlo (MC) code a constraint on the maximum number of bonds each patch can be involved in. In this way, only one particle can occupy the bonding volume, contributing u_0 to the potential energy. A double-stranded DNA with palindromic sticky overhangs at its extremities would be a possible experimental realization of such a particle. A careful choice of the strand sequences would make it possible to fulfil the single bond per patch condition and to control both the extension and the flexibility (i.e. the bonding volume) of the sticky ends.

Temperature is measured in units of the potential depth (i.e. $k_B/u_0 = 1$, where k_B is Boltzmann constant). The choice of a simple square-well interaction model to describe the association process between different particles is particularly convenient from a theoretical point of view. It allows for a clear definition of bonding and a clear separation of the bonding free energy in its energetic and entropic contributions, being unambiguously related to the depth of the well and to the bonding volume, respectively. In addition, the persistence length of the polymer is T-independent, due to the flatness of the attractive potential.

MC simulations in the NVT ensemble, with $N = 5000$ have been performed for five different densities $\rho = 0.000005, 0.00007, 0.0005, 0.007, 0.02$ and five different temperatures $T = 0.05, 0.06, 0.07, 0.08, 0.09$. The KF model parameters were set to $\Delta = 0.3$ and $\cos\theta_{\max} = 0.7077$, corresponding to a bonding volume $v_b = \frac{4}{3}\pi[(\sigma+\Delta)^3 - \sigma^3](1-\cos\theta_{\max})^2$. During equilibration, especially at very low ρ and T, we have implemented the Aggregation Volume Bias move proposed in Ref. [146], specialized to patchy particles. These moves are very effective at low densities, since they allow for a fast equilibration of the bonded-non bonded chemical equilibrium. Such a move is described in depth in Chapter 3.

2.3 Wertheim theory

The first-order Wertheim thermodynamic perturbation theory (W-TPT) [13, 130, 131] provides an expression for the free energy of associating liquids. The Helmholtz free energy is written as a sum of the HS reference free energy F_{HS}

plus a bond contribution F_{bond}, which derives by a summation over certain classes of relevant graphs in the Mayer expansion [13]. In the sum, closed loops graphs are neglected. The fundamental assumption of W-TPT is that the conditions of steric incompatibilities are satisfied: (i) no site can be engaged in more than one bond; (ii) no pair of particles can be double bonded. These steric incompatibilities are satisfied in the present model thanks to the location of the two sites and the chosen numerical algorithm which prevents multiple bonding.

The inclusion of the contribution of rings to the free energy is done like in [128, 134, 135]: the bonding free energy per particle of an homogeneous system is expressed as,

$$\beta F_{bond}/N = \ln X_0 + 1 - 2X_A + \frac{X_A^2}{X_0} - f_c - f_r, \tag{2.1}$$

where X_0 is the fraction of particles that have no patches bonded ("free" monomers), X_A is the fraction of unbonded patches, and f_c and f_r are, respectively, the contribution of chains and of rings to the free energy. Under the assumptions of W-TPT, f_c is,

$$f_c = 2\rho\Delta_{AA}X_A^2, \tag{2.2}$$

where ρ is the number density, and

$$\Delta_{AA} = \int g_{HS}(\vec{r})\left[\exp(\beta u_0) - 1\right] d\vec{r} = v_b\left[\exp(\beta u_0) - 1\right]. \tag{2.3}$$

In the previous expression, the pair correlation function of the reference fluid has been approximated by $g_{HS} = 1$, since only very low densities will be studied. The ring contribution is generalized to include the possibility of formation of rings of every size, and thus,

$$f_r = \frac{G_0}{\rho} \tag{2.4}$$

with,

$$G_i = \sum_n n^i \left(2\rho\Delta_{AA}X_0\right)^n W_n. \tag{2.5}$$

This expression corresponds to the summation of closed loop graphs, and is calculated assuming that, consistently with the usual W-TPT, the bonds of a ring are independent. W_n is proportional to the number of configurations of a single ring with n bonds. In previous works [128,134,135], W_n has been approximated

using the exact expression for freely jointed chains [147]. Here we will use an expression based on simulation results (see sections IV and V-A). It is also worth noticing that G_i is the i^{th} moment of the ring size distribution. The formation of a ring of size n requires n particles with the two bonding sites available to form bonds: the density of such particles is ρX_0. Each of these particles can bond in two different orientations (thus the factor 2 in Eq. (2.5)) and the probability to form a bond, once the available sites of the two particles are chosen, is Δ_{AA}.

The bonding free energy presented in Eq. (2.1) is then a function of X_0 and X_A and, at equilibrium, must be stationary with respect to variations of these densities. Minimizing Eq. (2.1) with respect to X_A and X_0, gives, respectively,

$$-1+\frac{X_A}{X_0}-2\rho\Delta_{AA}X_A=0, \tag{2.6}$$

and,

$$1-\frac{X_A^2}{X_0}-\frac{G_1}{\rho}=0. \tag{2.7}$$

These are the laws of mass action that establish the number of bonds and the number of rings in equilibrium, for a given set of thermodynamic variables (ρ, T). By replacing Eq. (2.6) in Eq. (2.1), one obtains for the equilibrium bonding free energy per unit volume,

$$\beta F_{bond}/V=\rho\ln X_0+\rho-\rho X_A-G_0. \tag{2.8}$$

Notice that when there are no rings (i.e. $G_i=0$ in the previous equations), one recovers the usual expressions obtained when only chain aggregates are considered (see e.g. [57, 142]).

Thermodynamic calculations are straightforward, once an approximation for W_n is adopted. For fixed (ρ, T) the fractions X_0 and X_A are calculated using Eqs. (2.6) and (2.7) and the bonding free energy follows from Eq. (2.8).

2.4 Equivalence with an ideal gas of chains and rings description

The free energy per unit volume of an ideal mixture of chains and rings can be written as

$$\beta F_{id.cl.}/V=\sum_n \rho_n^c(\ln\rho_n^c-1-\ln\tilde{q}_n^c)+\sum_n \rho_n^r(\ln\rho_n^r-1-\ln\tilde{q}_n^r) \tag{2.9}$$

where ρ_n^c is the density of chains of size n, ρ_n^r is the density of rings of size n, and $\tilde{q}_n^c$ and $\tilde{q}_n^r$ are the partition functions of a ring and of a chain of size n, divided by the volume V. The chemical equilibrium between chains and rings gives,

$$\frac{\partial \beta F/V}{\partial \rho_n^c} = \frac{\partial \beta F/V}{\partial \rho_n^r} = n\beta\mu \tag{2.10}$$

where μ is the chemical potential. Consistent with the ideal-gas hypothesis, the number density of monomers (i.e. particles with no bonds) ρ_1^c is related to μ by $e^{\beta\mu} = \rho_1^c$. Eqs (2.9) and (2.10) give origin to the following expressions,

$$\frac{\rho_n^c}{\tilde{q}_n^c} = \frac{\rho_n^r}{\tilde{q}_n^r} = \exp(n\beta\mu) = (\rho_1^c)^n, \tag{2.11}$$

from which the relation between the chain and rings size distribution and partition functions follows,

$$\rho_n^r = \rho_n^c \frac{\tilde{q}_n^r}{\tilde{q}_n^c}. \tag{2.12}$$

This shows that in the ideal-gas (of cluster) limit the ring size distribution is not an independent variable, but it is proportional to the chain size distribution times the ratio between the ring and chain partition functions.

The chain partition function can be written as

$$\tilde{q}_n^c = \exp\left[-(n-1)\beta\mathcal{F}_b\right] \tag{2.13}$$

in terms of the free energy of a bond, $\mathcal{F}_b$. This expression corresponds to assuming that each of the $n-1$ bonds lowers the free energy by $\mathcal{F}_b$. The ring partition function depends on the geometry of the model for short rings, while it reaches a well established asymptotic value for large n values, dictated by the self-similar nature of the growth process. More specifically, in the dilute limit,

$$\tilde{q}_n^r \propto \frac{v_b}{nR_{ee}^3(n)} \exp\left(-n\beta\mathcal{F}_b\right) \tag{2.14}$$

where v_b is the bonding volume and $R_{ee}(n)$ is the mean end-to-end distance of a chain of size n. Compared to linear polymers, rings are thus stabilized by the presence of n bonds (one more than a chain). On the other hand, the number of configurations allowing for ring formation are proportional to the ratio $v_b/R_{ee}^3(n)$. The additional n dependence arises from the number of ways a ring can open to form a chain [125, 126, 148]. Assuming that chains can be modelled with a self-avoiding walk process for large n (i.e. $R_{ee}(n) \sim n^{\nu_{SAW}}$,

where $\nu_{SAW} = 0.588$ is the self-avoiding walk exponent), one can thus write

$$\tilde{q}_n^r = \alpha(n) n^{-(3\nu_{SAW}+1)} \exp\left(-n\beta\mathcal{F}_b\right) \tag{2.15}$$

where $\alpha(n)$ describes the model dependent short-n behaviour and approaches a constant value for large n, when the self-similar nature of the ring is reached. The evaluation of $\alpha(n)$ for the chosen model is discussed in the following (section V-A).

Expression (2.9) can be simplified by using Eq. (2.11) to obtain

$$\beta F_{id,cl.}/V = \rho \ln \rho_1^c - \sum_n (\rho_n^c + \rho_n^r). \tag{2.16}$$

Once approximations for $\mathcal{F}_b$ and $\alpha(n)$ are adopted, ρ_n^c and ρ_n^r can be expressed as a function of ρ_1^c via Eq. (2.12) and the latter can be calculated at a given (ρ, T) imposing the normalization condition,

$$\sum_n n(\rho_n^c + \rho_n^r) = \rho. \tag{2.17}$$

In order to demonstrate the equivalence between the ideal cluster of chains and rings description and that obtained using Wertheim theory (with the ideal gas as the reference system), we note that the quantities X_0 and X_A can be expressed in term of rings and chains distributions as

$$\rho X_A = \sum_n \rho_n^c, \tag{2.18}$$

and

$$\rho X_0 = \rho_1^c. \tag{2.19}$$

The bonding free energy $\beta\mathcal{F}_b$ can be related to the Mayer function of the bond via $\exp(-\beta\mathcal{F}_b) = 2\Delta_{AA}$. Finally, the number of ring configurations W_n can be identified with

$$W_n = \alpha(n) n^{-(3\nu_{SAW}+1)}. \tag{2.20}$$

Via this set of equivalences, $\rho_n^r = (\rho_1^c)^n \tilde{q}_n^r$ can be written as $(2\Delta_{AA}\rho\ X_0)^n W_n$ and Eq. (2.16) transforms to

$$\beta F_{id,cl.}/V = \rho \ln(\rho X_0) - \rho X_A - \sum_n (2\Delta_{AA}\rho\ X_0)^n W_n. \tag{2.21}$$

By defining the ideal gas free energy $F_{id} \equiv \rho \ln(\rho) - \rho$ and exploiting the defini-

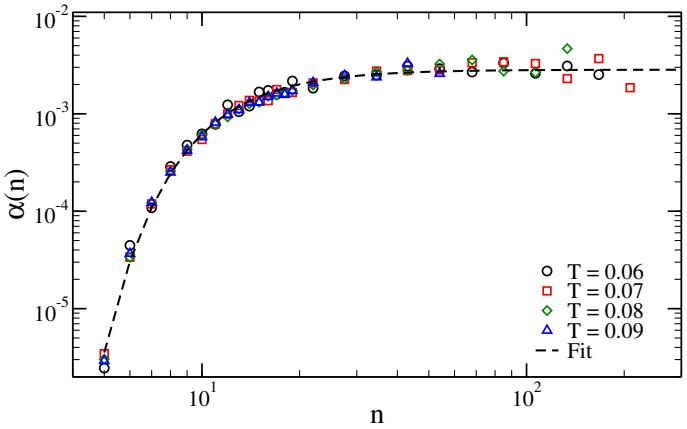

Figure 2.1: Numerical evaluation of $\alpha(n)$, Eq. (2.23), at four different temperatures and $\rho = 0.007$. The line is the fit function Eq. (2.24).

tion of G_0 in Eq. (2.5) and of $\beta F_{bond}/V$ in Eq. (2.8), Eq. (2.21) becomes

$$\beta F_{id,cl.}/V = \beta(F_{bond} + F_{id})/V \tag{2.22}$$

proving that the free energy of an ideal gas of chains and rings can be written as the sum of the Wertheim bonding free energy and of the reference ideal gas free energy.

2.5 Results

2.5.1 Estimate of the ring partition function

The evaluation of the free energy of a system of chains and rings requires the knowledge of the function $\alpha(n)$ introduced in Eq. (2.15). As discussed in Section 2.4, such function encodes all non-universal model properties and needs to be evaluated independently. For models (like the present one) in which the persistent length does not depend on T, the function $\alpha(n)$ does not depend on T and can thus be estimated at any (low density) state point, by evaluating the ratio between ρ_n^r and ρ_n^c. Indeed, from Eq. (2.12) and Eq. (2.15),

$$\alpha(n) \equiv \frac{\rho_n^r}{\rho_n^c} n^{3\nu_{SAW}+1} \exp\left(-\beta \mathcal{F}_b\right) \tag{2.23}$$

The choice of the $n^{3\nu_{SAW}+1}$ scaling, strictly correct only in the dilute limit [126], is consistent with the ideal-gas of clusters approximation.

Fig. 2.1 shows $\alpha(n)$ evaluated at four different state points. The asymptotic

value is reached only for chains longer that $30-50$ particles. The minimum ring size is five. The numerical $\alpha(n)$ can be represented by the arbitrary fit function

$$\ln(2v_b\alpha(n)) = -5.86595\left[1+35.4697\cdot\exp\left(-\frac{\ln n}{0.468144}\right)\right], \tag{2.24}$$

and such expression has been used to evaluate theoretically all system properties.

2.5.2 Chain and ring distribution functions

Fig. 2.2 shows the chain and the ring distribution functions for different T values at $\rho = 0.02$. The chain distributions are exponential, with a characteristic decay length that increases on cooling. The ring distributions are non-monotonic and increase their width and amplitude on cooling, signalling the progressive increasing concentration of rings. The position of the peak in the ring distribution functions, whose presence is due to the existence of a preferred cluster size determined by the interplay between energy and entropy, is model dependent and can be tuned by changing the angular width of the patches.

Fig. 2.3 and 2.4 show the average size of chains M^c and rings M^r, defined respectively as $M^c = \sum n\rho_n^c / \sum \rho_n^c$ and $M^r = \sum n\rho_n^r / \sum \rho_n^r$, as a function of inverse T and ρ. Both M^c and M^r increase monotonically upon cooling the system and increasing density, but the effect of changing ρ and T is much more dramatic on M^c. Indeed, M^c spans three order of magnitude in the studied range of parameters, whereas M^r changes of a mere factor of 5 going from $T = 0.09$, $\rho = 0.000005$ to $T = 0.06$, $\rho = 0.02$.

The theoretical predictions shown in these figures are calculated using,

$$M^c = \frac{1-G_1/\rho}{X_A}, \tag{2.25}$$

and

$$M^r = \frac{G_1}{G_0}. \tag{2.26}$$

Given the low densities studied (that make Wertheim theory almost exact) and the numerical approximation adopted for the partition function of rings, the very good agreement with the simulation results is expected. In Figures 2.3 and 2.4 the result obtained when rings are neglected is also shown. One can conclude that the effect of rings in the chain size distribution becomes important at low ρ and T. For fixed low ρ, the decrease in T increases this influence. On the

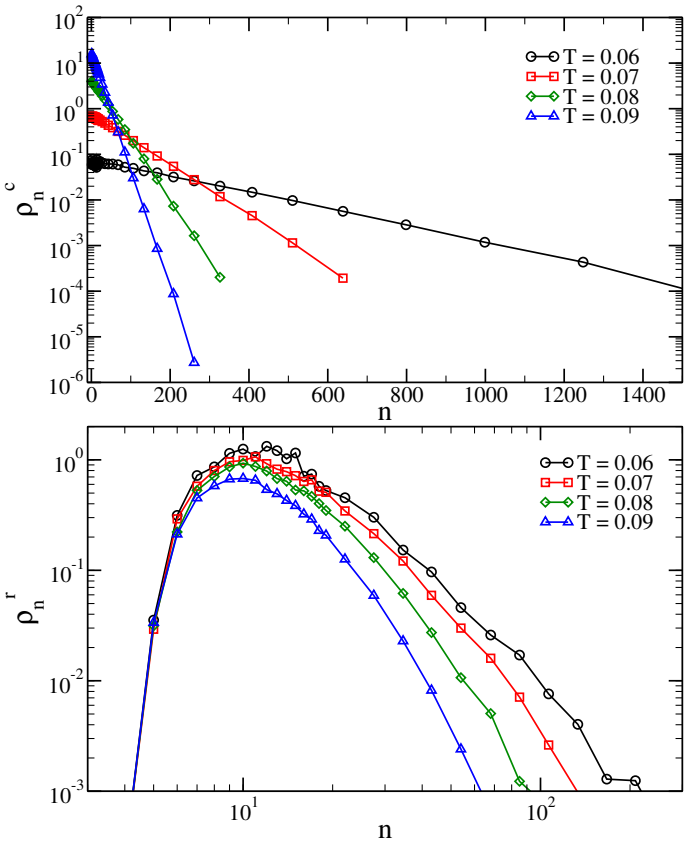

Figure 2.2: Chain (top) and ring (bottom) size distribution functions at $\rho = 0.02$ and different temperatures T.

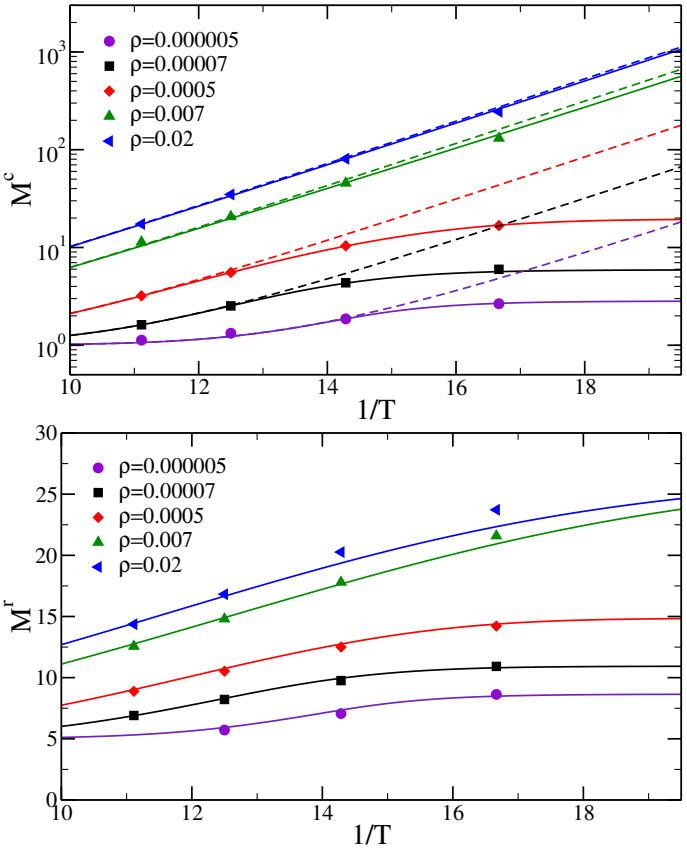

Figure 2.3: Average cluster size of chains (top) and rings (bottom) as a function of the inverse temperature for different ρ. Symbols are numerical results; solid lines are the theoretical predictions of Eqs. (2.25) and (2.26); dashed lines are the theoretical predictions when rings are excluded from Wertheim theory.

other hand, for fixed low T, this effect is important in a range of intermediate ρ; this range increases for lower T.

2.5.3 The onset temperature

We have seen in Section 2.4 that in the limit of non-interacting clusters, the ratio between ring and chain distributions is controlled by the ratio of the ring and chain partition functions. Such ratio has a maximum at a specific ring size and approaches zero both at smaller sizes, due to the difficulty to form rings of sizes comparable to the persistence length, and at large sizes, where the entropy cost of localizing chain-ends controls aggregation. Fig. 2.5 shows the ratio $\frac{\rho_n^r}{\rho_n^c}$ for the present model, at five different T and $\rho = 0.02$. All the curves retain the same shape, peaking at $n_{max}^c \approx 10 - 11$, but their amplitude strongly depends on T,

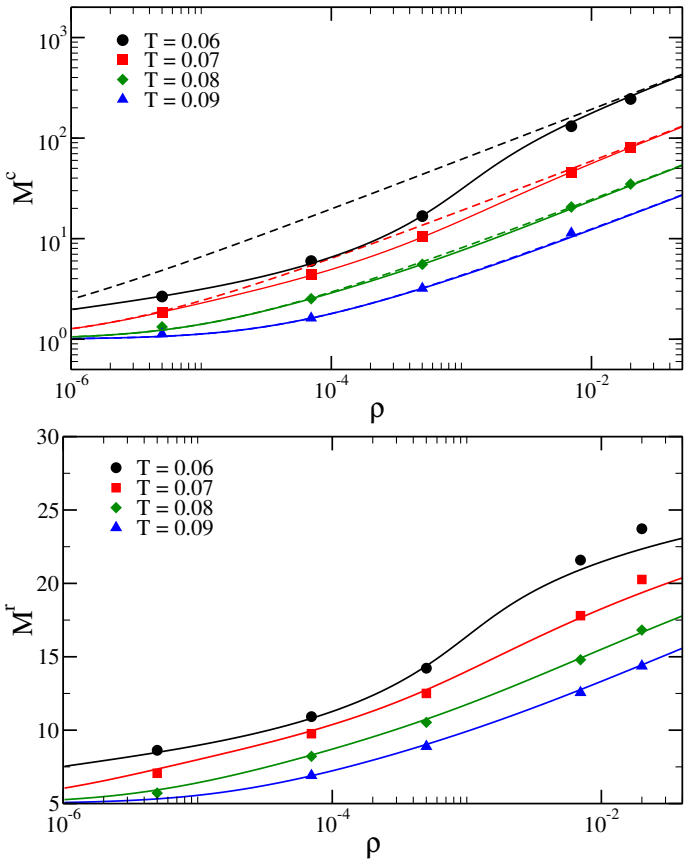

Figure 2.4: Average cluster size of chains (top) and rings (bottom) as a function of ρ for different T. Symbols and lines as in Fig. 2.3.

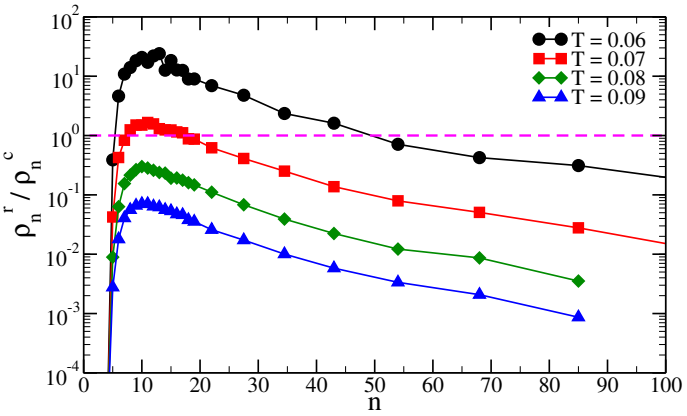

Figure 2.5: Ratio between ρ_n^r and ρ_n^c at four different T and $\rho = 0.02$. The dashed line indicates $\rho_n^r/\rho_n^c = 1$, to visualize the range in n for which, at a given $T > T_{onset}$, rings become more probable than chains.

via the factor $\exp \beta\mathcal{F}_b$. The T dependence provides a way of defining an onset temperature for rings (T_{onset}), for example as the T at which $\frac{\rho_{nmax}^r}{\rho_{nmax}^c} = 1$, as

$$\exp\left(\frac{\mathcal{F}_b}{k_B T_{onset}}\right) = \alpha(n_{max}) n_{max}^{-(3\nu_{SAW}+1)}. \tag{2.27}$$

For the present model, $T_{onset} \approx 0.075$.

2.5.4 Potential energy

Fig. 2.6 shows the potential energy per particle U/N (proportional to the number of bonds) for different T and ρ. The energy decreases in a small interval of T, approaching a value $-u_0$, which describes the condition in which all possible bonds are formed, at the smallest investigated T. The figure also shows the Wertheim theoretical predictions, both including and excluding the possibility of ring formation. At large T, rings are negligible and both theories provide the same predictions, but around and below the onset temperature, the presence of rings significantly modifies the T-dependence of U, giving rise to a faster approach to the ground state energy. The extended Wertheim theory captures extremely well the T and ρ dependence of the energy.

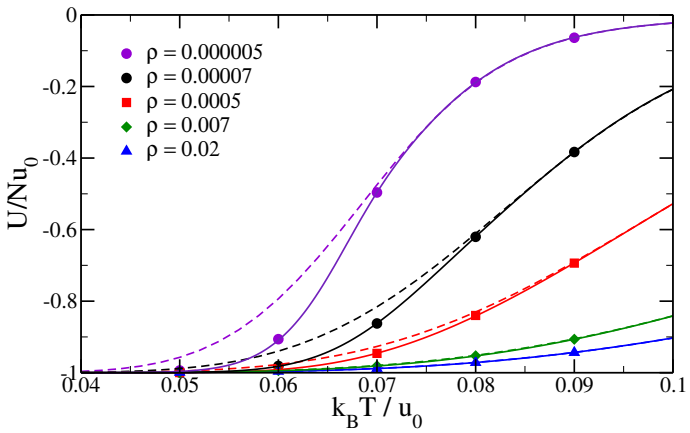

Figure 2.6: Potential energy (proportional to the number of bonds) per particle as a function of T for several ρ values. Symbols are MC results. Lines are theoretical predictions for the case in which rings are either excluded or included in Wertheim theory.

2.5.5 Fraction of particles in rings

To globally assess the behaviour of the system, we evaluate the T and ρ dependence of the fraction of particles in rings, defined as

$$f_{pr} = \frac{\sum_n n\rho_n^r}{\rho} = \frac{G_1}{\rho} \tag{2.28}$$

The results are shown in Fig. 2.7. The quantity f_{pr} shows a non-monotonic behaviour as a function of ρ, decreasing both for small and large densities. The non-monotonic behaviour of f_{pr} is a consequence of the non-monotonic n dependence of the ratio $\frac{\rho_n^r}{\rho_n^c}$, which is strongly peaked around n_{max}. Along a constant T path, the concentration of rings is expected to have a maximum when the average size of the aggregates is of the order of n_{max} (see Fig. 2.4(b)). Hence, f_{pr} decreases when ρ is so small that only very small clusters are present as well as when ρ is so large that the average chain length is much longer than n_{max}. Indeed, on increasing ρ, the number of bonds in the system increases (see Fig. 2.6), favouring the formation of chains longer than n_{max}.

Fig. 2.7 shows also the theoretical evaluation of f_{pr} based on the Wertheim theory, with no fitting parameters. As for the potential energy, the theory very accurately captures the ρ and T dependence of f_{pr}.

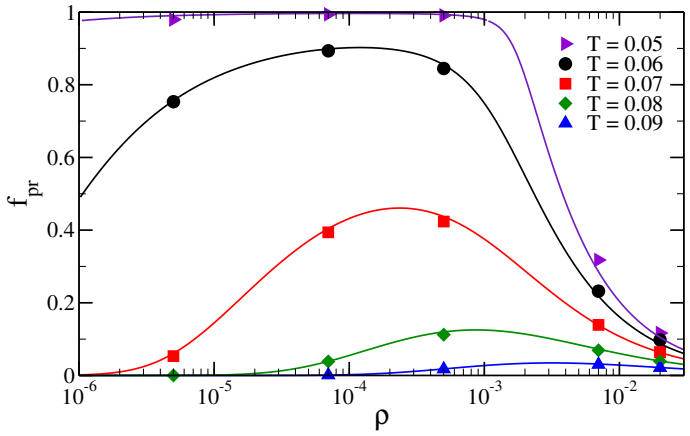

Figure 2.7: Fraction of particles in rings f_{pr} evaluated from MC simulations (symbols), and with Wertheim theory (lines).

2.6 Conclusions

Recent studies [133] have shown that an accurate modelling of chaining in patchy particle systems in dilute conditions can be achieved via the Wertheim thermodynamic perturbation theory, without resorting to any fitting parameter. In this Chapter we have extended this approach to the case in which chaining competes with the formation of ring structures, once more comparing numerical results with predictions based on the Wertheim thermodynamic perturbation theory, augmented to include the description of close loops of bonded particles [128, 134, 135]. While the Wertheim theory provides an expression for the chain partition function, valid in the limit of persistent chains, the ring partition functions is one additional input to the theory. We have resorted to a numerical evaluation of such a quantity, consistent with the expected limits — based on the statistic of self-avoiding walks — for large ring sized. The resulting model-dependent expression properly describes all ring sizes at low ρ, incorporating the non-universal properties of the KF-model.

We find that the augmented Wertheim theory properly captures the ρ and T dependence of the aggregation phenomenon. It is able to predict with extreme accuracy the non-monotonic behaviour of the fraction of particles in rings on increasing ρ at low T. Indeed, the system evolves from a collection of very short chains at extremely small densities, to a state in which almost of particles are in rings, and back to a state in which long polymers dominate again for large densities. The region where rings dominate is confined at low T (where the presence of an additional bond energetically stabilizes rings as compared to a

chain of corresponding size) and at low densities (where the average cluster size is limited). The ρ and T dependence of the number of bonds (i.e., the energy) or of the average cluster size of chains and rings is also predicted with very high accuracy.

The ability of Wertheim theory to also model the delicate equilibrium between chains and rings, together with the previously explored ability of the Wertheim theory to model the thermodynamics of chaining and branching, suggests that it will be possible to combine the two approaches and derive an accurate theory for systems in which rings, chains and branched structures coexist. It is particularly important to find out under which conditions the branching-induced gas-liquid critical point survives in the presence of ring formation. It has indeed been shown [39] that low-valence patchy particles are characterized by a branching driven critical point at rather low densities, i.e. in the same region where rings can be dominant. The scenario in which rings, chains and branched structures coexist and compete has been recently observed in low T simulations of dipolar hard sphere fluids, in both two [149] and three [62] dimensions, one of the canonical models for self-assembly. It has been suggested that the observed increase in ring structures could possibly provide an example of phase separation suppressed by self-assembly [61]. Theories developed to describe the gas-liquid phase separation in these systems mainly focus on the delicate balance between chaining and branching [55–57]. Since the presence of rings alters this balance, a new theoretical framework which also takes into account rings is needed. The remarkably good agreement between numerical results and theoretical predictions based on the augmented Werheim reported in this Chapter strongly suggests to tackle the low-T DHS behaviour by further extending the theory to incorporate branching. Such a study is under way.

Chapter 3

Between bivalent and trivalent systems: dipolar fluids

3.1 Introduction

The dipolar hard sphere (DHS) model is of paramount importance in the physics of disordered systems [59, 60, 150–153]. In the framework of limited-valence colloids which is the topic of this Thesis, DHSs can be regarded as having a temperature-dependent valence, which, as will be discussed in the following, tends to 2 as temperature and density are decreased.

The DHS model consists of a point dipole embedded in the center of a hard sphere, and it is the simplest model which incorporates anisotropic long-range interactions. Notwithstanding its simplicity, the DHS model is still the subject of thorough investigations, aimed in particular at understanding its low-temperature and low-density behaviour. From a theoretical point of view, the main difficulty arises from the fact that at low-T and low-ρ the phase behaviour of the model is determined by the competition between condensation and self-assembly. The condensation of dipolar particles, similar to the usual gas-liquid phase transition, was first predicted by de Gennes and Pincus [58] after observing that the spherically averaged interaction between dipoles is attractive. On the other hand, dipolar particles self-assemble in linear and branched structures with the dipoles aligned in the (energetically favourable) nose-to-tail

geometry [60, 154–157]. Indeed the first computer simulation studies, although plagued by equilibration issues, provided evidence of this extended nose-to-tail chaining and found contradictory evidence of a phase transition [59, 152, 158]. The anisotropic aggregation of particles which results from the self-assembly process was long thought to be the mechanism responsible for the suppression of isotropic condensation and the absence of a phase transition between disordered states. This belief was completely reverted by theoretical studies [55] which showed that self-assembly alone is capable of sustaining a phase transition. The transition is hierarchical and involves a subtle balance between chain ends, which provide an effective repulsion, and *Y-shaped* chain junctions, which provide an effective attraction. The resulting competition leads to two coexisting phases with different topology: a "gas" of chains and a network-like "liquid" rich in junctions. The possibility of such a phase transition, which depends crucially on the number density of topological defects and their scaling with density results in a peculiar re-entrant phase diagram, in which the density ρ of the liquid phase approaches the vanishing ρ of the gas phase on cooling. The essential properties of the *topological* phase transition can be studied both by mean-field theories [55] and thermodynamic perturbation theories [56, 57], showing that criticality arises only if the ratio between the energy cost of junction formation over the energy cost of chain-end formation is within a specific range of values. Whether such energy ratio in the DHS potential is consistent with the possibility of observing a topological phase transition is still unknown. A recent thermodynamic perturbation theory [159], in which chain association is explicitly accounted for, suggests instead the existence of a standard gas-liquid coexistence.

The low-T and low-ρ properties of the zero-field DHS model in three dimensions ($3D$) are hard to grasp not only from a theoretical point of view, but also pose significant challenges to both experimental and computational studies. From an experimental point of view one is faced with the difficulty of generating single crystal magnetic cores of suitable size and without any residual isotropic attraction (which would artificially promote the condensation of the dipoles). From the computational point of view the major challenges come from the expensive long range electrostatic calculations and from the long relaxation times of the structures (chains and networks) into which the dipoles self-assemble. Most of the simulation studies have thus been plagued by equilibration issues or severe finite size effects. These difficulties have inspired a lot of works to study model systems which are easier to solve numerically and that converge

asymptotically to the DHS as a function of some parameter [151,160–162]. Based on extrapolation of DHS plus an attractive isotropic Yukawa component [151], binary mixtures of apolar and dipolar hard-spheres [161] or charged dumbbells [160, 162] the location of the putative gas-liquid critical point has been restricted to the window $\rho \lesssim 0.1$ and $T < 0.16$.

Another line of research concentrated on the mapping between patchy particle models and DHS [143]. A first attempt, which focused on the study of temperature-dependent valence models [56, 57], successfully located the topological phase transition and found a very good agreement with theoretical predictions. Another example, patchy particles which self-assemble in rings and chains which then compete at low densities and temperatures, is provided in Chapter 2.

In this Chapter, we use highly efficient simulation techniques and extended computational resources to address the question of whether a gas-liquid phase transition exists in DHS. These techniques allow us to access in equilibrium the temperature region where theories and previous numerical attempts had predicted a gas-liquid coexistence [151, 159–162]. We definitively prove that no sign of critical behaviour is observed. The analysis of the low-T data reveals that the reason for the absence of the topological phase transition in the investigated phase diagram region is not related to a high energy cost of junction formation, as in one of the expected theoretical scenarios, but to the breakdown of the mean-field approximations at low-T and low-ρ. These results open up a new scenario for DHS, where new theoretical modelling is needed in order to explain its phase behaviour.

As a first step in this endeavour, in the second part of the Chapter we carefully study the structural properties of the DHS fluid at low-T and low-ρ in order to generate information that can be used to provide appropriate parameters which can help mapping DHS into patchy particles models [143] that can be analytically solved within the Wertheim formalism [130, 131]. In Chapter 2 we make a step forward in this direction. We provide detailed structural information as well as a study of the aggregate topologies to provide accurate data for refined theoretical modelling. In particular, we focus on chains and rings, which are the most common structures found in low-T, low-ρ DHS fluids [140, 149]. The interplay between these structures has been the center of numerous studies. Small-angle neutron scattering [138, 163, 164] and TEM [165] experiments have shown that the structure factor at small wavelengths displays a power-law behaviour due to the aggregation of particles in finite-size and

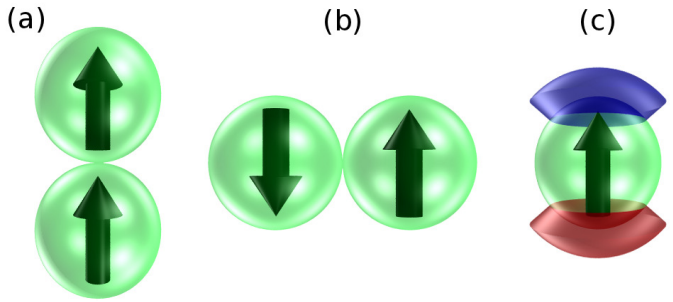

Figure 3.1: (a) Nose-to-tail contact configuration corresponding to an absolute minimum in the interaction energy of $u = -2$ (in natural units). The arrows represent the point dipole embedded in the hard spheres. (b) Anti-parallel contact configuration corresponding to a relative minimum of $u = -1$. (c) A single DHS with northern (blue) and southern (red) virtual bonding regions highlighted.

long-lived clusters. Many numerical and theoretical approaches were devoted to interpreting the structural signatures of low-T ferrofluids [136, 166–170]. Simulations in quasi two dimensions suggested that at very low T and ρ the system is mainly composed of isolated rings. Indeed, it has been theoretically shown that the ground state in quasi-2D systems is a single ideal ring for sizes larger than 4 [171]. As density is increased the average size of rings increases until they break into open chains [139] due to entropic effects. As density is further increased the chains start branching until they form a percolated network [149].

We find that the DHS model at low T exhibits remarkable and unusual behaviours like a very low density percolation locus and a stabilization of rings over chains, resulting in a depletion in the number of chain ends. We speculate that this excess number of rings is a possible mechanism for suppressing the gas-liquid phase separation.

In Section 3.2 we present the model and the state-of-the-art numerical techniques employed to tackle this long-standing open problem. Section 3.3 shows results on phase behaviour and structural properties of the DHS fluid at low densities and temperatures. In Section 3.4 we summarise our findings and outline possible extensions.

3.2 Methods

3.2.1 Model and computational details

The pair interaction potential between two dipolar hard spheres i and j is

$$u(i,j) = u_{HS}(r_{ij}) + \frac{\boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j - 3(\boldsymbol{\mu}_i \cdot \hat{\mathbf{r}}_{ij})(\boldsymbol{\mu}_j \cdot \hat{\mathbf{r}}_{ij})}{r_{ij}^3} \tag{3.1}$$

where $\mathbf{r}_{ij}$ is the vector connecting the centres of particles i and j, $r_{ij} = |\mathbf{r}_{ij}|$, $u_{HS}(r_{ij})$ is the hard sphere potential and $\boldsymbol{\mu}_i$ is the dipole moment of particle i. In the following, the Boltzmann constant $k_B = 1$, $\beta = 1/T$, lengths are measured in units of the particles diameter σ and energy in units of μ^2/σ^3. In these units, the most energetically favourable configuration is the nose-to-tail contact geometry, with a energy contribution of $u = -2$ (Fig. 3.1(a)). This absolute minimum is responsible for the characteristic chain-like structures which become prominent at low T. A relative minimum arises from the side-by-side anti-parallel geometry, corresponding to a pair interaction energy of $u = -1$ (Fig. 3.1(b)).

The long range nature of the dipolar interaction and the clustering process hinder the possibility of carrying out equilibrium simulations in the low T ($T \lesssim 0.2$), low ρ region via conventional methods. Indeed, when thermal fluctuations become of the order of one tenth of the head-tail energy, a significant chaining process starts to take place. In order to explore the low-T and low-ρ region in equilibrium we implement special Monte Carlo state-of-the-art moves and techniques specifically aimed at the DHS system.

Long range dipolar interactions are taken into account using Ewald sums with conducting boundary conditions [172]. In order to improve performance we use a mesh-based cubic spline interpolation for cosines and sines calculations, resulting in a $\sim 30\%$ speed-up. Such approximation yields a precision in the energy evaluation of 10^{-6}.

Aggregation-volume bias algorithm for dipolar particles

One of the biggest difficulties in the simulation of self-assembly processes is the fact that the structures which are formed are located in a very small region of the phase space, and are thus difficult to sample. Moreover, once these structures are formed, the breaking of bonds becomes a rare event, which prevents sampling of independent configurations. An efficient way to overcome

these difficulties is to facilitate the formation and breakage of bonds via biased Monte Carlo moves [172]. In this work we adapt the aggregation-volume-bias (AVB) algorithm [146] which, given the definition of the region around each particle where bonds are stronger, accelerates both bond formation and bond breakage by moving particles inside and outside of these regions.

For DHS the most favourable bonding configurations correspond to the head-tail geometry, as shown in Fig. 3.1(a), and we thus define two virtual bonding regions (BR) on the poles of each particle. A bonding move consists in selecting the bonding region of a particle (say for example the north pole), and then moving the complementary pole (south pole) of another particle inside this region. The reverse move consists instead in moving a particle out of the bonding region to a random location in the sample. Following the Kern-Frenkel idea [144], the bonding regions are shaped as truncated cones (Fig. 3.1(c)) of angular width $\theta = 0.873\ rad$ ($\cos(\theta) = 0.64$) and range $\delta = 0.4\sigma$. Two particles i and j are considered virtually bonded if their relative distance is smaller than $\sigma + \delta$ and if $\boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j > \cos\theta$, corresponding to a volume $V_{AVB} = \frac{4}{3}\pi\left[(\sigma+\delta)^3 - \sigma^3\right]\frac{1-\cos\theta}{2} = 1.3\,\sigma^3$. We define N_i as the number of particles which are in the BR of i. More precisely, two AVB moves are introduced [146]:

1. With probability p_{AVB} we choose a random particle i and then we choose, randomly, another particle j which is in the BR of i. We move particle j out of the BR of particle i, inserting it with a random orientation. If no particles are found in the BR of i, the move is *a priori* rejected. We accept the move with probability

$$p_{in\to out} = \frac{1-p_{AVB}}{p_{AVB}}\frac{N_i(4\pi V - V_{AVB})}{(N-N_i)V_{AVB}}e^{-\beta\Delta U}.$$

2. With probability $1 - p_{AVB}$ we choose a random particle i and a random particle j which is not inside the BR of i. We move particle j inside the BR of particle i with a proper orientation to guarantee virtual bonding. We accept the move with probability

$$p_{out\to in} = \frac{p_{AVB}}{1-p_{AVB}}\frac{(N-N_i-1)V_{AVB}}{(N_i+1)(4\pi V - V_{AVB})}e^{-\beta\Delta U}.$$

The value of p_{AVB} can be used to tune the relative acceptance probabilities for these two moves. A value of $p_{AVB} = 0.5$ was adopted throughout this work. We

have thoroughly tested the AVB algorithm for models of associating particles where chaining and branching are dominant [56, 57]. We found that the AVB moves allow for equilibration down to temperatures $k_B T/u_b \simeq 0.060$, where u_b is the typical bonding energy. This maps into $T \approx 0.12$ for the DHS case, where $u_b \approx 2$. Note that the implemented AVB algorithm only acts on the process of forming and breaking chain structures. The process of breaking/forming branching points is equilibrated by standard Monte Carlo roto-translational moves, i.e. moves in which a random displacement and/or rotation is applied to each particle. Indeed, branched configurations are characterized by u_b values smaller than the nose-to-tail ones and, in the context of associating particles models [56, 57], the AVB algorithm does not result in any significant speedup when encoded also for branching interactions. We fix a 1:1 frequency ratio between roto-translations and biased moves. In the following, we define a MC step as N attempts to perform a rotation/translation or an AVB move (where N is the number of particles in the system). A measure of the efficiency of the AVB method will be provided in Fig. 3.2, where the bond-bond autocorrelation function is reported.

Biased insertion/deletion moves

Since in the low-T, low-ρ DHS fluid most particles are part of a long-living polymer network, the energy required to extract one particle is very high and regular grand-canonical MC (GCMC) moves have a very low acceptance ratio. In order to overcome this limitation, we adopt the technique introduced by Caillol [150] (and adapted for grand canonical MC by Ganzenmüller *et al.* [151]) to speed-up equilibration and sampling. The value of the local electric field $\mathbf{E} \equiv \mathbf{E}(\mathbf{r})$ is computed, where $\mathbf{r}$ is either the trial inserting position for insertions or the position of the particle to be removed for deletions. Since $\boldsymbol{\mu} \cdot \mathbf{E} = \mu E \cos\theta$ is the interaction energy with the field, the Boltzmann distribution of the dipole orientation is given by

$$f(\cos\theta) = \frac{\beta\mu E \exp(\beta\mu E \cos\theta)}{2\sinh(\beta\mu E)}. \tag{3.2}$$

For insertions, a biased random angle between the dipole and the field can be readily obtained from Eq. (3.2) from a generated random number $R \in [0, 1]$ via

$$\cos\theta = \frac{1}{\beta\mu E} \log\left[2R\sinh(\beta\mu E) + \exp(-\beta\mu E)\right]. \tag{3.3}$$

The acceptance probabilities for such orientationally biased insertions ($N \to N+1$) and deletions ($N \to N-1$) become, respectively,

$$p_{N\to N+1} = \min\left[1, \frac{0.5}{f(\cos\theta)} \frac{zV\exp(-\beta\Delta U)}{N+1}\right] \tag{3.4}$$

$$p_{N\to N-1} = \min\left[1, \frac{f(\cos\theta)}{0.5} \frac{N\exp(-\beta\Delta U)}{zV}\right] \tag{3.5}$$

where z is the activity at which the simulation is performed. The factor 0.5 accounts for the value that the function $f(\cos\theta)$ takes in the ideal-gas case [151].

In GCMC simulations, the frequency of insertions/deletions, AVB moves and regular rotations/translations is, respectively, 1:50:50.

Successive Umbrella sampling

Finally, to uniformly sample the particle density of states (despite the large free energy difference between low and high density phases at fixed activity), and to effectively parallelise our simulations, we perform successive umbrella sampling (SUS) simulations [173]. With this method, the region to be explored is partitioned in overlapping windows of ΔN particles. Each window is then sampled with GCMC simulations with appropriate boundary conditions [174], providing a speed-up proportional to the number of windows explored in parallel (200 in our case). The SUS method allows us to obtain the distribution of density fluctuations $P(\rho)$ at fixed activity $z = e^{\beta\mu}$ (where μ is the chemical potential). We then evaluate $P(\rho)$ at different z and nearby T by means of histogram reweighting techniques [175].

3.2.2 Bond definition and cluster classification

To quantify the connectivity properties of the system a definition of a bond between a pair of particles is required. Given the continuous nature of the dipole-dipole interaction, bonding can not be unambiguously defined. In the past, criteria based on cut-off distances [149, 151] or pair interaction energy thresholds [59, 151, 166, 168, 176–178] have been proposed. Since sharp energy thresholds may lead to an underestimation of the branching, which is a key factor in topological phase transitions, and since the position r_b of the first minimum of the $g(r)$ is not very sensible to T or ρ (see Section 3.3.5), we combine the two approaches and define in the following particles i and j as bonded

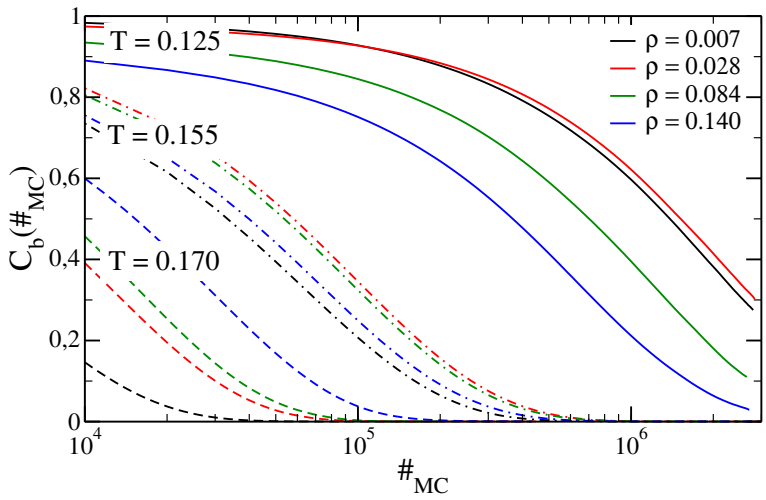

Figure 3.2: Bond-bond autocorrelation function C_b as a function of the number of MC steps $\#_{MC}$ for four different densities and $T = 0.125$ (full lines), $T = 0.155$ (dashed-dotted lines) and $T = 0.170$ (dashed lines).

if $r_{ij} < r_b = 1.3$ and $u(i,j) < 0$. This criterion is somewhat similar to the "entropic" criterion proposed by Holm *et al.* [168], with the difference being that the criterion employed here takes into account also the role of the branching due to anti-parallel geometry, which is not considered by the latter. Since this type of branching occurs only at intermediate ($\rho \gtrsim 0.05$) densities, the two criteria are equivalent in the low-ρ limit.

To provide evidence that the length of our simulations is sufficiently long to probe equilibrium states, we evaluate the bond-bond autocorrelation function $C_b(\#_{MC})$. This quantity is defined as the probability that a bond existing at the beginning also exists after $\#_{MC}$ steps, quantifying how fast the network topology rearranges. Fig. 3.2 shows $C_b(\#_{MC})$ for three temperatures: $T = 0.125, 0.155$ and 0.170. It is interesting to observe that, at very low T, the system loses memory of its initial state faster at higher ρ. This is consistent with the fact that, at such T, the lowest energy geometry is the one in which a particle is in the core of an isolated infinitely long chain. We also note that the shape of the correlation function can be properly modelled via a weakly stretched exponential (stretching parameter ≈ 0.9), reflecting the different local bonding geometries.

Employing the previous bonding criterion, we classify all clusters according to their bonding topology in three groups:

- **chains**: clusters containing two ends (i.e. particles with just one bonded

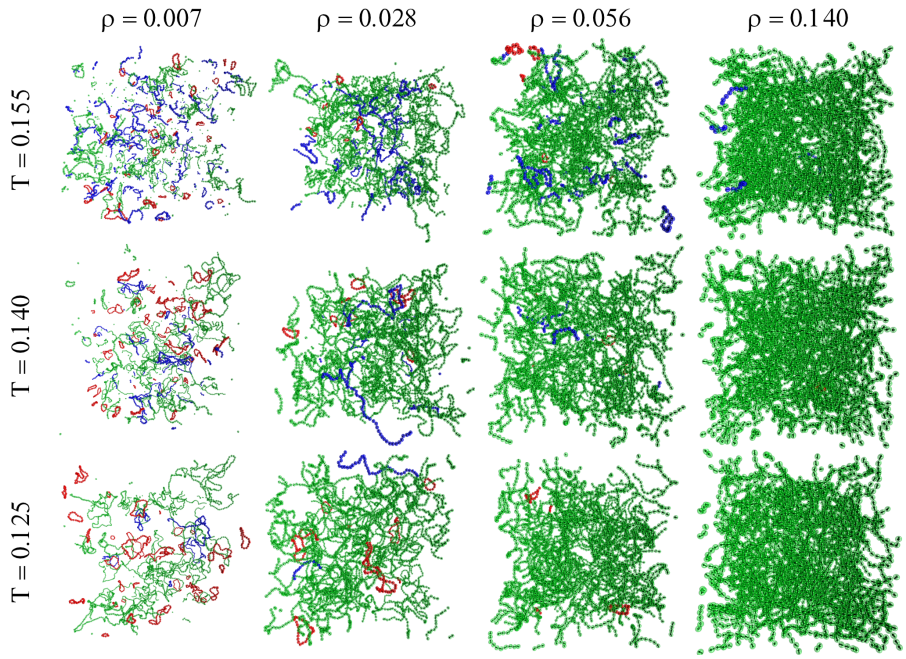

Figure 3.3: Typical snapshots of equilibrium configurations of DHS at $T = 0.125, 0.140, 0.155$ and $\rho = 0.007, 0.028, 0.056, 0.140$. The different cluster topologies are depicted with different colors: chains (blue), rings (red) and branched structures (green).

neighbour) connected by particles with only two bonded neighbours;

- **rings**: clusters containing only particles with two bonded neighbours;
- **branched structures**: clusters containing at least one particle with more than two bonded neighbours.

3.3 Results

To provide a pictorial representation of the studied systems, Fig. 3.3 shows snapshots of equilibrium configurations of DHS systems at three different temperatures and for the whole range of studied densities. Two trends are already clearly distinguishable. First, as ρ increases, both chains (in blue in the picture) and rings (in red) progressively merge together to form branched structures (in green). Secondly, a decrease in T is accompanied with an increase in the number of rings, at the expenses of the chain structures.

In the following we study in detail the thermodynamic and structural properties at all these state points.

3.3.1 Distribution of density fluctuations

We have investigated three different T, namely $T = 0.125$, 0.140 and 0.150, all below the lowest estimate $T = 0.153$ at which the critical point is predicted to be located [151, 159, 161, 162, 179]. We investigate a system of up to $N = 1000$ particles in a box size of $L_0 = 19.26$ via windows, overlapping by one particle, of width $\Delta N = 6$. This corresponds to the density region $\rho \in [0, 0.14]$.

Fig. 3.4 shows the main result of this Chapter, $P(\rho)$ for several values of z and T. On decreasing temperature, the noise level increases, signalling the difficulty of properly sampling configuration space, despite the length of the simulation (extending to ten months of computation for each of the 200 Xeon 5550 CPU cores). For all T, the shape of $P(\rho)$ excludes the presence of a gas-liquid coexistence down to $T = 0.125$. Two minor peaks are also found, both located at very low densities. A peak at $N \sim 10 - 15$ (corresponding to $\rho \sim 0.0015 - 0.002$) is observed at all T. One further peak around $N \approx 20 - 25$ appears at $T = 0.14$ and becomes prominent at the lowest T, as shown in the inset of Fig. 3.4(c). To identify the nature of these two peaks we perform a finite size study of the low density region at $T = 0.140$. We run simulations at different box sizes L, chosen in such a way that $L^3 = k \cdot L_0^3$ with k integer. In absence of finite size effects, if a box of volume V is split in k small parts of volume V_0, then the probability of having N particles in the larger box depends on the probabilities of finding $0, 1, \ldots, N$ particles in the smaller box, and thus an expression that directly connects the distribution of density fluctuations at a given volume $V = L^3$ $P(N; V)$ to the original $P(N; V_0 = L_0^3)$ can be derived. For example, if $k = 2$ and $N = 2$, then $P(2; V) = 2P(2; V_0)P(0; V_0) + 2P(1; V_0)^2$. The generalised expression reads

$$P(N, kV_0) = \sum_{i_1} \sum_{i_2} \ldots \sum_{i_k} \delta(N - i_1 - i_2 - .... - i_k) P(i_1, V_0) P(i_2, V_0) \ldots P(i_k, V_0). \tag{3.6}$$

Fig. 3.5 shows that the second peak is present only for the smallest box, and hence it is definitely a finite size effect. The inset, showing the difference occurring for $N \gtrsim 25$ in the rescaled distributions of density fluctuations, confirms this hypothesis. The first peak at $N \sim 10 - 15$, on the other hand, is a feature of the system and not a finite size effect since it is always present and its location does not scale with the box size. The presence of constant N peaks is typically associated with self-assembly processes of aggregates with a preferential size,

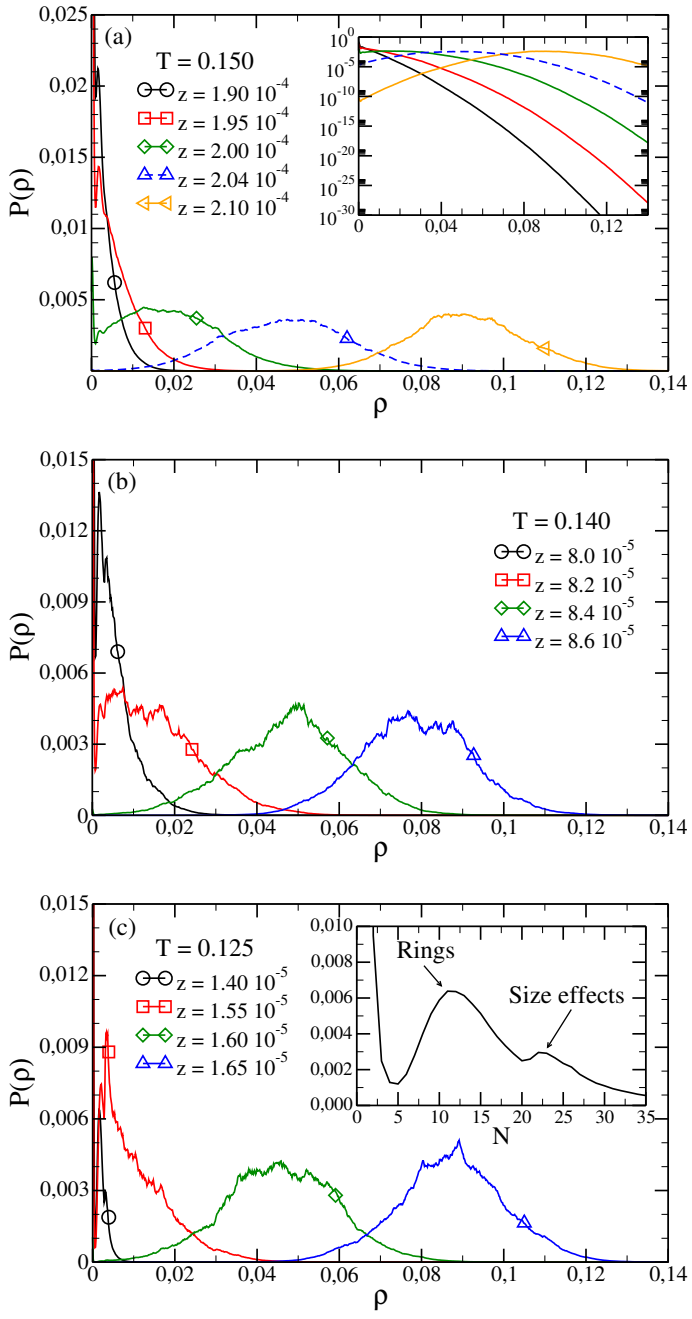

Figure 3.4: Distribution of density fluctuations $P(\rho)$ for different activities z at (a) $T = 0.150$, (b) $T = 0.140$ and (c) $T = 0.125$. Curves are obtained via histogram reweighting in z, starting from the value of z at which the simulations were performed: (a) $z = 2.04 \cdot 10^{-4}$ (dashed line), (b) $z = 9.00 \cdot 10^{-5}$ and (c) $z = 1.535 \cdot 10^{-5}$. Points are added on each curve in order to increase b/w readability. Inset: (a) lin-log representation of $P(\rho)$. (c) $P(N)$ for $z = 1.4 \cdot 10^{-5}$ in the range $N \in [0, 35]$.

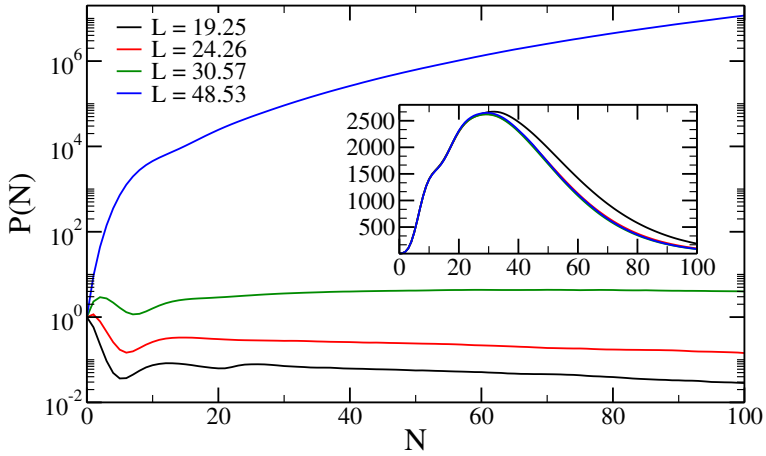

Figure 3.5: $P(N)$ for $N \in [0, 100]$ at $T = 0.140$ for different box sizes. In the inset all the $P(N)$ are rescaled on top of each other using Eq. (3.6) and reweighted so that the main peak indicating the density of the phase in the thermodynamic limit is located at $N \approx 30$. The difference between the rescaled $L = 19.25$ curve and the other $P(N)$ for $N \gtrsim 30$ signals that the small peak at $N \approx 25$ found in the former is due to finite size effects.

which does not depend on the system volume [180, 181]. The second peak reflects the presence of linear chains which percolate via boundary conditions. Compared to chains of the same size, they are energetically stabilized by the presence of one additional bond.

As the DHS system does not exhibit any sign of gas-liquid criticality in the window $0.125 < T < 0.150$ and $0 < \rho < 0.14$, i.e. even well below the region where such critical point was predicted to be located, we are interested in the mechanism responsible such a suppression of the phase separation. We speculate that the absence of gas-liquid criticality is related to the fact that the low ρ non percolating fluid, being very rich in rings, is in a state in which all particles are close to their ground state energy and which has a larger translational entropy compared to the percolating phase. In order to back up this claim, in the next sections we perform large-scale NVT simulations of DHS systems composed of $N = 5000$ particles at six different densities ($\rho\sigma^3 = 0.007$, 0.028, 0.056, 0.084, 0.114, 0.140) and four different temperatures ($T = 0.125$, 0.140, 0.155, 0.170).

We note that the box sizes employed in the simulations (whose values range from 33 to 90, depending on the density) are about one-order of magnitude larger than the average chain persistence length (see Section 3.3.6). In addition, we check that chains never percolate. Moreover, when a percolated network is

present, the mesh size of its structure is always significantly smaller than the box length.

3.3.2 Cluster size distributions

To clarify the origin of the self-assembly peak we investigate in detail the structure of the fluid at low T, analysing NVT configurations. The number of particles employed ($N = 5000$) and the large box sizes, beside improving the quality of the data, suppress any finite size effects associated with percolating chains. Based on the bonding criterion adopted, we partition particles into chains, rings or branched clusters [59,149,151,177], according to their topology. We find that clusters with $s \lesssim 40$ are mostly rings or chains, i.e. only a few junctions are present in small clusters. Beyond a certain ρ, the system is always percolating, that is, more than 50% of its equilibrium configurations contains a spanning (infinite) cluster. The percolation ρ becomes as small as $\rho \approx 0.01$ when $T = 0.125$, as shown in Section 3.3.6. The ρ dependence of the cluster size distribution $n(s)$ is shown in Fig. 3.6(a). A peak at $N \approx 10 - 15$, i.e. at the same location as the one observed in $P(\rho)$, is present at all T, confirming that such a peak arises from the preferential self-assembly of the particles in particular clusters. Separating $n(s)$ in its ring ($n_r(s)$, Fig. 3.6(b)), chain ($n_c(s)$, Fig. 3.6(c)), and branched components, allows us to identify the rings as the structures responsible for the peak. Indeed, $n_r(s)$ is non-monotonic, shows a tail which becomes progressively extended on cooling and is peaked at $N \approx 10 - 15$. $n_c(s)$ is monotonic and decays exponentially only for large s, suggesting that the free energy cost of adding a particle to a chain becomes independent of the chain length only for $s \gtrsim 20$. The slope of the exponential tail decreases on cooling, signalling the progressive increase of the average chain length [59,177].

At the lowest T rings of size $10 \lesssim s \lesssim 100$ become more probable than chains of the same size and the total number of particles in rings N_r (inset of Fig. 3.6(b)) becomes larger than the number of particles in chains N_c (inset of Fig. 3.6(c)). Such an increase in N_r on lowering T and ρ offers a possible hint on why the critical phenomenon is not observed. Indeed, rings are characterized by a small net total dipole moment, resulting in a small effective ring-ring interaction. The low density DHS thus progressively turn into a fluid of weakly interacting aggregates, providing an example of phase separation suppressed by self-assembly [182,183].

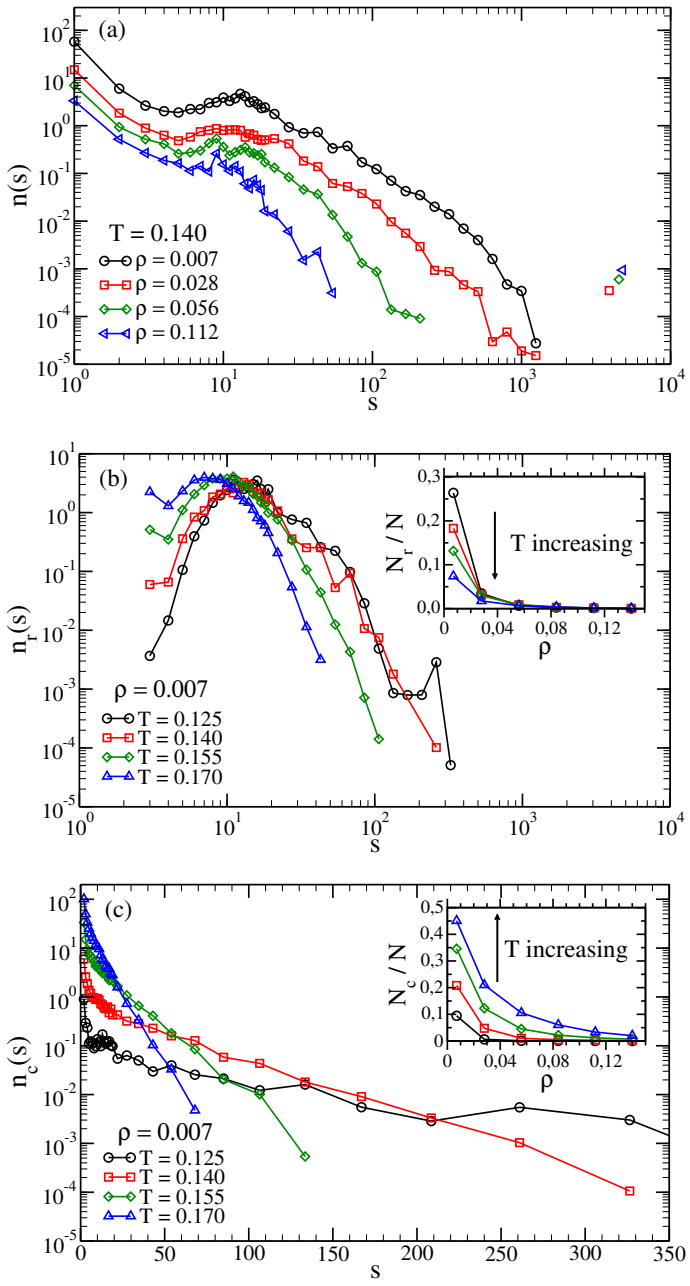

Figure 3.6: (a) Number of clusters $n(s)$ of size s at $T = 0.140$ and different ρ in a system of $N = 5000$ particles ($\sum_s sn(s) = N$). Clusters with $s > 3000$ (disconnected points) are percolating. Number of rings $n_r(s)$ (b) and of chains $n_c(s)$ (c) of size s at $\rho = 0.007$ at different T. The inset shows the fraction of particles in rings N_r/N (b) and in chains N_c/N(c) as a function of ρ. Note the inverted order with T of N_c/N and N_r/N as well as the exponential tail only for large s in $n_c(s)$. Lines are guides for the eye.

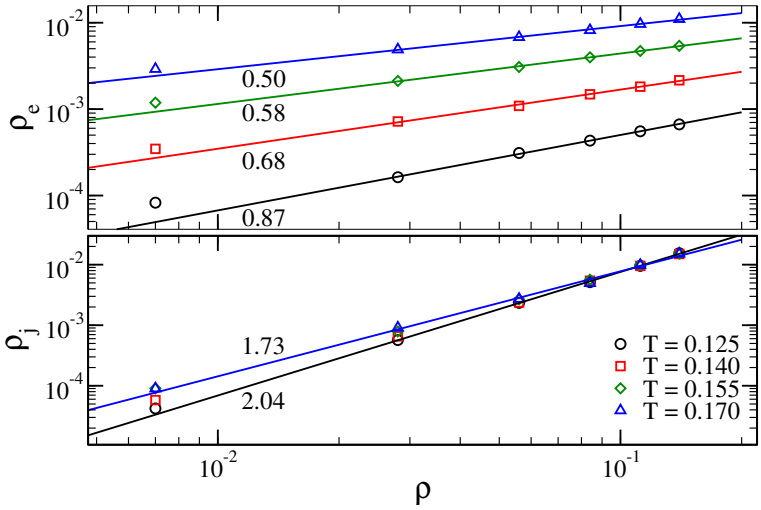

Figure 3.7: Density of ends ρ_e and of junctions ρ_j as a function of ρ at all the studied T. Lines are power-law fits restricted to $\rho > 0.007$. The value of the best-fit exponent is indicated beside the lines. A significant T-dependence is observed. Note that a drastic change in the bonding distance criterion results only in a slight change in the value of the junction exponent and in a 20% change in the end exponent, but both exponents retain their T dependence.

3.3.3 Chain-ends and Y-shaped junctions

To test the topological phase transition hypothesis we examine the ρ dependence of the concentration of chain ends ρ_e (particles with just one bonded neighbour) and of junctions ρ_j (particles with three bonded neighbours). In the mean-field TS theory both ρ_e and ρ_j follow a power law in ρ with exponents $1/2$ and $3/2$, respectively. These exponents play a major role, controlling the ρ dependence of the system free energy and thus dictating the origin of the phase separation. Fig. 3.7 shows ρ_e and ρ_j for all the studied state points. We first note that only for $\rho > 0.007$, do ρ_e and ρ_j follow a power-law, but with a T-dependent exponent. Moreover, the exponents appear to approach the values assumed in the TS theory only at the highest T studied. This strongly suggests that the competition between topological defects is not a viable mechanism for sustaining a critical point in DHS.

3.3.4 Potential energy

Next we compute the configurational potential energy per particle U. Fig. 3.8 shows U as a function of ρ for all studied T. The energy shows a very weak ρ dependence, especially on lowering T, a clear signature of the onset of a

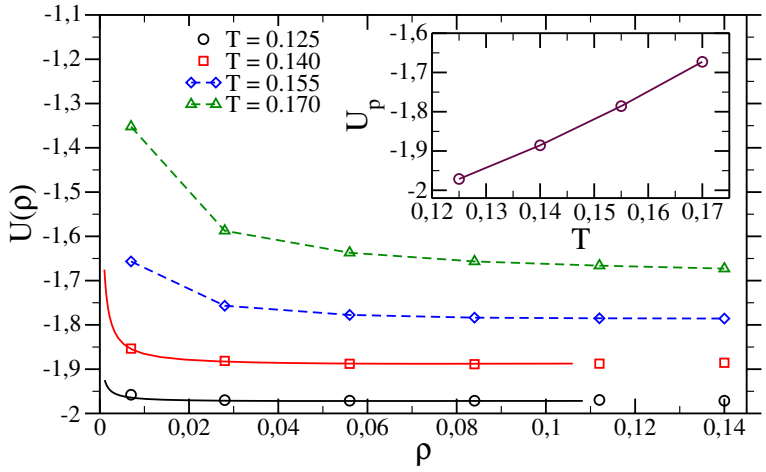

Figure 3.8: Potential energy per particle $U(\rho)$ as a function of ρ for different T. Points are NVT results, solid lines are μVT results. Dashed lines are guide for the eye. Inset: height of the plateau U_p for $T = 0.125$, 0.140 and 0.155.

self-assembly process [155, 184]. Indeed, self-assembly is characterized by the formation in the fluid phase of well defined structures (clusters), in which particles have a characteristic energy. When increasing density results only in an increase in the number of such structures, then the energy does not depend significantly on density.

To provide hints on the origin of the energy saturation with ρ, we examine the energy of clusters of different size. As we will discuss in more details in the following, at the lowest ρ most of the particles are in chains and rings. On increasing ρ, the number of particles in chains and rings drops in favour of extended branched structures, which percolate beyond a critical ρ. Here we focus primarily on chains and rings. The intra-cluster energy of chains and rings in the context of the DHS model has been previously addressed [149]. An expression for the ground state energy per particle $e_c(s)$ in a chain of size s has been proposed, based on the assumption that chaining produces a rescaling by a factor α (to take into account beyond-nearest neighbours interactions), of the dimer head-to-tail ground state energy $-2\mu^2/\sigma^3$. According to this assumption, each bond contributes an energy $\alpha(2\mu^2/\sigma^3)$. Since there are $s-1$ bonds in a chain, $e_c(s, T = 0) = \alpha(2\mu^2/\sigma^3)\frac{s-1}{s}$. This ground state expression has been generalised at finite T [149] by retaining the s dependence as:

$$e_c(s) = -e_0^c + \frac{e_1^c}{s}. \tag{3.7}$$

Here e_0^c is the bond energy for infinite long chains and e_1^c accounts for finite size

effects.

A similar approach can be applied to rings. Since the angle between the dipole moments of two nearest-neighbouring particles in a ring of size s is $2\pi/s$, the ring energy at $T = 0$ can be written as $\alpha(2\mu^2/\sigma^3)\left[3 + \cos\left(\frac{2\pi}{s}\right)\right]/4$, whose finite T generalisation reads [149]

$$e_r(s) = -e_0^r + \frac{e_1^r}{s^2}. \tag{3.8}$$

All the coefficients, namely e_0^c, e_0^r, e_1^c, and e_1^c, are expected to depend only on T.

Fig. 3.9 shows $e_r(s)$ and $e_c(s)$ as a function of the cluster size for a system at $T = 0.125$ and $\rho = 0.007$ and the corresponding fit according to Eq. (3.7) and (3.8). The fit properly represents the s dependence of the particle energy. The asymptotic value e_0^c representing the average bond energy in large chains, is reached only when $s \gtrsim 100$. The asymptotic value e_0^r, referring to rings, is instead reached when $s \gtrsim 50$, as expected on the basis of the s^{-2} dependence in $e_r(s)$. e_0^c and e_0^r are almost ρ independent and the fit provides the same value e_0 within statistical error. The T dependence of e_0 is presented in the inset of Fig. 3.9. e_0 decreases with T, possibly due to the reduced vibrational contribution to the potential energy. To confirm the vibrational nature of the T dependence of e_0 we evaluate the inherent structures (IS) [185,186], by minimizing the energy of a set of equilibrium configurations at $\rho = 0.007$ and $T = 0.125$. The functional form describing the s dependence of the chains and rings remains identical. The average bond energy in the IS configuration is -2.33, to be compared to the value of the energy per particle in the *close-packed* antiferromagnetic configuration, equal to -2.56 [58]. The inset also shows that the T dependence of e_0, including the $T = 0$ inherent structure value, can be modelled assuming a weak anharmonicity as $e_0(T) = -2.34 + 2.5k_BT + 3.84(k_BT)^2$.

The best-fit values of the remaining fitting parameters (e_1^c and e_1^r) are shown in Fig. 3.10. Such values are very similar to the ones found in $2D$ DHS [149].

3.3.5 *g(r)* and *S(q)*

To quantify the structural changes of the DHS fluid at different T and ρ, we compute the radial distribution function $g(r)$. The $g(r)$ provides information on the relative distance between particles in the system. Fig. 3.11(a) shows the T effect on $g(r)$ at densities $\rho = 0.007$ and $\rho = 0.140$ (inset). The $g(r)$ is well structured even at the highest investigated T. Upon lowering T, minima

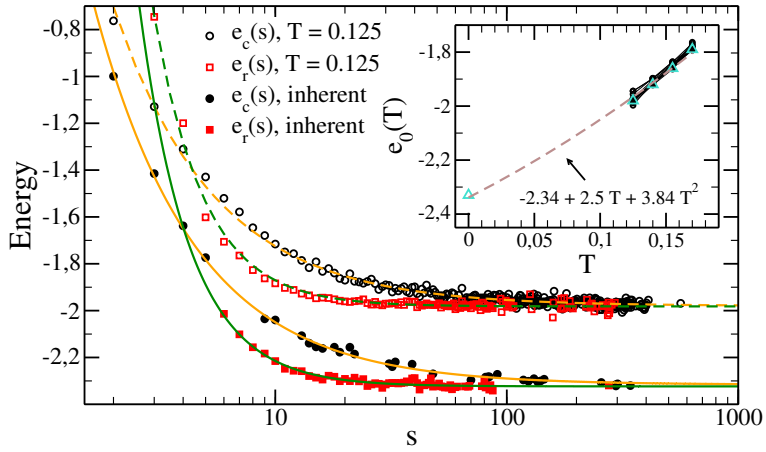

Figure 3.9: Potential energy of particles in chains (open circles) and rings (open squares) of size s at $\rho = 0.007$ and $T = 0.125$ and for the inherent structure (full symbols). Lines are fit to the simulation data according to Eqs. (3.7) (orange dashed and full lines) and (3.8) (green dashed and full lines). Inset: e_0 at $\rho = 0.007$ for all studied T and for the inherent structure (triangles). To provide evidence of the ρ independence of e_0 we also show the corresponding value for all investigated densities (small black circles). The brown curve is a quadratic fit to the triangle points in which the harmonic linear term $2.5T$ has been fixed.

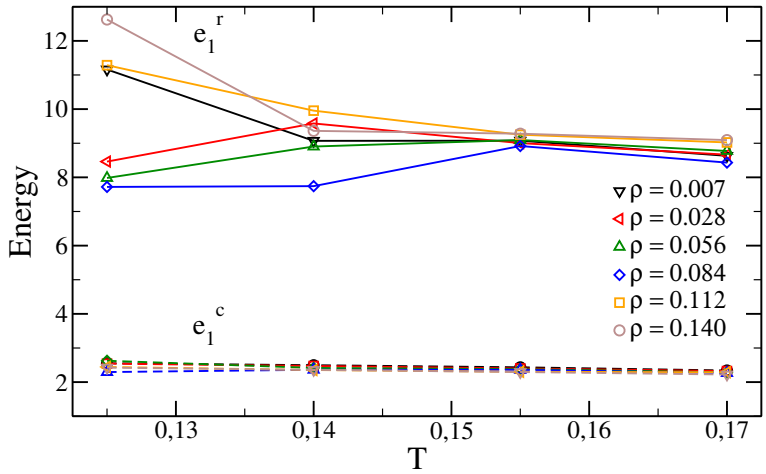

Figure 3.10: Fit parameters e_1^c (dashed lines) and e_1^r (full lines) for all the state points.

deepens and the height and sharpness of the peaks increases. At large r, $g(r)$ approaches one from above, reflecting the presence of spatial inhomogeneities, intrinsic to the cluster structure of the system. Similarly, the radial position of the third and of the following peaks increases. The first two phenomena highlight the increasing bond localization which takes place in this T region. The shift of the secondary peaks towards values of r which are multiple of σ suggests that chains become straighter on cooling, or, in other words, the persistence length of the chains increases. It is interesting to observe that the first peak becomes more and more resolved on cooling, and that the amplitude of the first minimum approaches zero. Under these conditions, the bonding criterion based on the relative distance becomes precise.

At higher ρ (inset of Fig. 3.11(a)), the system always contains a percolating structure (see Section 3.3.6) and the $g(r)$ becomes more similar to the one observed in simple fluids. Now at large r, $g(r)$ oscillates with periodicity σ around one. The shape of the peaks is still asymmetric, reflecting the preferential one-dimensional growth of the equilibrium aggregates.

In order to clarify the effect of ρ on the structure, we study the number of neighbours $n(r)$ within a sphere of radius r centred on an arbitrary particle, defined as

$$n(r) = 4\pi\rho \int_0^r r^2 g(r')dr'. \tag{3.9}$$

The result is shown in the inset of Fig. 3.11(b) for $T = 0.125$. For low and intermediate ρ, a plateau develops for $1 < r < 2$ in which $n(r) \simeq 2$, a clear evidence that, locally, particles are coordinated with only two neighbours, i.e. that the system associates mostly in chains and rings. At high ρ, the number of neighbours in the first shell increases, signalling the presence of branching and the formation of more complex structures.

The large value of the low q limit of the structure factor $S(q)$, calculated as $S(q) = \frac{1}{N}\left\langle \sum_{i=1}^{N}\sum_{j=0}^{N} e^{i\mathbf{q}\cdot(\mathbf{r_i}-\mathbf{r_j})}\right\rangle$ and presented in Fig. 3.12, provides another signal of the strong association occurring in the system. Indeed, in a system of independent (ideal gas) clusters, $S(q)$ reflects the properly averaged cluster form factor and it approaches the second moment of the mean cluster size when $q \to 0$. The most visible T effect can be found in the shape of the first peak, which becomes more and more asymmetric upon lowering T, an effect due to the presence of longer and longer chains [3,176]. At larger ρ (see Fig. 3.12(b)), interference between different clusters (which are now connected in a percolating network), significantly lowers the small-q value of $S(q)$. Interestingly, no

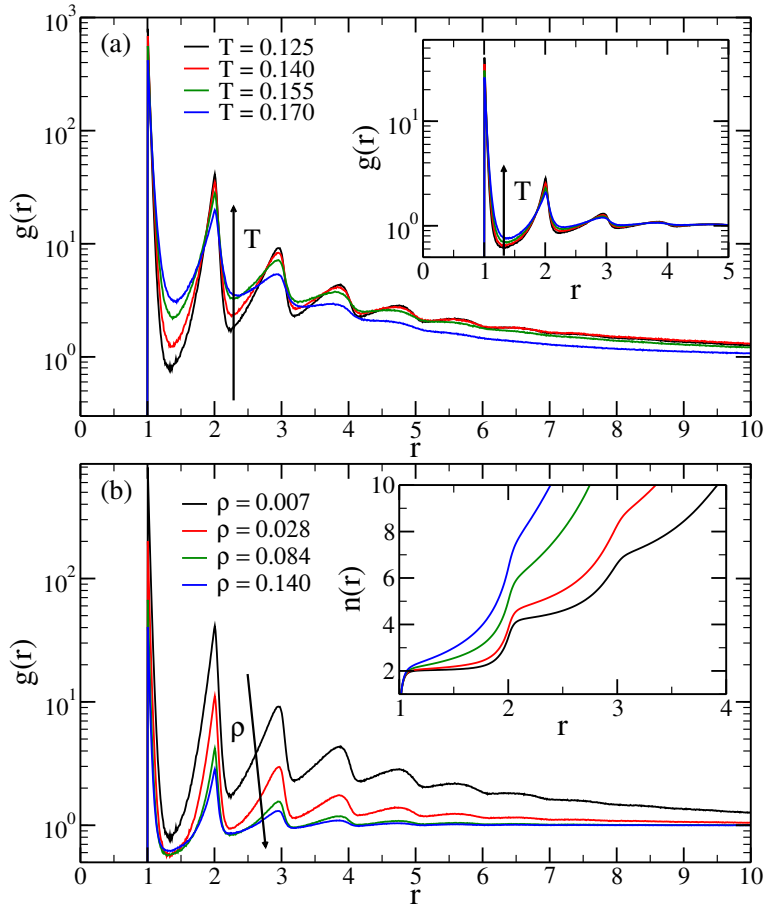

Figure 3.11: (a) $g(r)$ for all the studied temperatures at $\rho = 0.007$ (main panel) and $\rho = 0.140$ (inset). (b) $g(r)$ at fixed $T = 0.125$ for different densities. Inset: average number of neighbours $n(r)$ for the same state points.

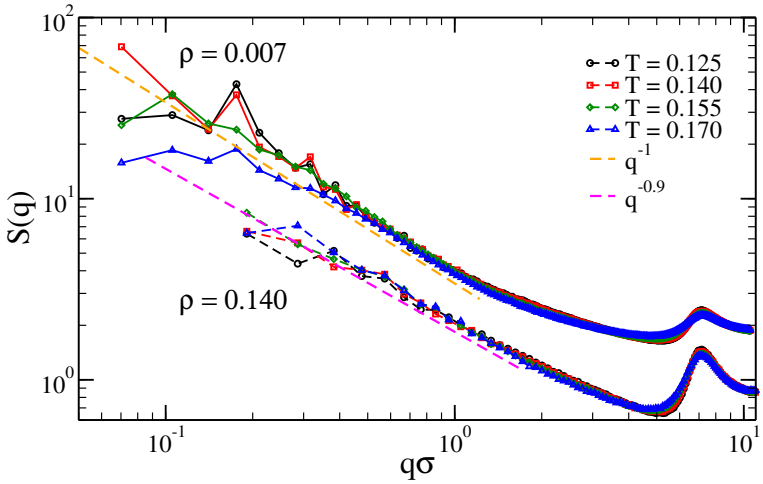

Figure 3.12: $S(q)$ for all the studied temperatures at $\rho = 0.007$ (full lines) and (b) $\rho = 0.140$ (dashed lines). The $\rho = 0.007$ data has been shifted upwards by adding 1. Dashed lines are power-law curves with exponents -1 (orange) and -0.9 (magenta).

significant T dependence is observed, suggesting that the system has reached its final structure and only minor changes take place. Such a property is characteristic of equilibrium gels [3,145]. The only effect of decreasing T is visible in the small increase in the height and asymmetry of the first peak.

$S(q)$ shows a clear change of "scaling" behaviours in different q-windows.. The low-q behaviour of the structure factor can often be related to the fractal dimension D of the aggregates via $S(q) \propto (q\sigma)^{-D}$. As Fig. 3.12 shows, the region where a power-law behaviour is clearly and unambiguously observed is less than one decade. Under these conditions, it is extremely hard and possibly misleading to extract exponents. A slope of the order of one, consistent with a strong signature from chains, is consistent with the scaling at small q in the region $0.3 \lesssim q \lesssim 1$ [170].

3.3.6 Connectivity properties and degree of polymerization

We now explore the global connectivity properties of the DHS model. In particular, we evaluate the percolation locus, separating percolating and non-percolating state points. Percolation (via physical interactions) is a pre-requisite for the onset of a second-order critical phenomenon [187]. Indeed, both in spherically interacting potentials [188], as well as in limited-valence patchy interactions [189], on cooling the percolation line is always encountered before the gas-liquid instability line. We define a state point as percolating if more

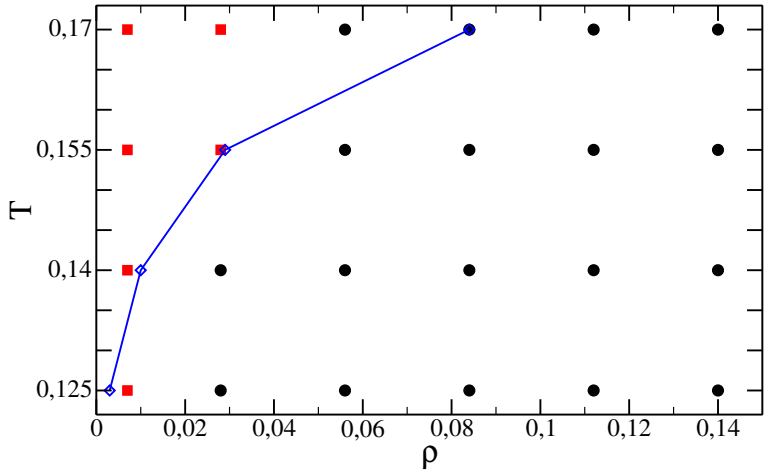

Figure 3.13: Percolating (black circles) and non percolating (red squares) state points. The percolation line passes near the $T = 0.125$, $\rho = 0.007$ state point, which has almost half of the configurations ($\sim 49\%$) containing a spanning cluster. The blue symbols are the state points where the degree of polymerization $\Phi \approx 0.99$.

than half of its configurations contains a spanning cluster, i.e. a cluster which is connected to itself via periodic boundary conditions. The results of this analysis are summarized in Fig. 3.13. The percolation threshold extends to very low ρ (as low as $\rho \approx 0.01$ for the lowest studied T in equilibrium). This result once more suggests that the system is composed of rather long chains connected via branching points.

Another interesting observable in the study of the self-assembly process is the so-called degree of polymerization Φ, commonly defined in self-assembly studies as the fraction of particles associated into clusters [133, 190] ($1 - \Phi$ is conversely the fraction of particles in monomers). In the T-ρ window explored in the present study, the degree of polymerization is always very close to one (we find that $\Phi > 0.95$ for every studied state point but $T = 0.170$, $\rho = 0.007$). To provide a reference for further study we report in Fig 3.13 the locus $\Phi = 0.99$.

3.3.7 Chains and rings at low densities

In this section we investigate the properties of chains and rings at low ρ ($\rho = 0.007$). We aim at providing accurate quantitative data on the structure of these clusters in the limit in which excluded volume cluster-cluster interactions are limited and the system can be in first approximation considered as a mixture of poly-disperse non-interacting clusters.

The fraction of particles involved in chains, rings and other branched structures at $\rho = 0.007$ and $\rho = 0.028$ is shown in Fig. 3.14. As shown by the figure, upon lowering T at fixed ρ, the fraction of particles in chains diminishes and the majority of the particles belongs to large branched structures. On further cooling, these particles would give rise to a percolating structure, despite the small average ρ. The decrease in the number of particles in chains on cooling has thus a double origin. On one hand, rings becomes more stable than chains (at comparable size) and indeed the number of particles in rings increases on cooling, and on the other hand, chains become longer and hence have a larger probability to branch. Despite this effect, even at the lowest T, a significant fraction of particles is still in chains and rings. This allows us to carry out an analysis of the properties of such structures, similar to what has been done in the past in the investigation of equilibrium polymerization in which ring formation is allowed [125, 126]. It is interesting to observe that the condition $n_r(T_x) = n_c(T_x)$ implicitly defines a cross-over temperature T_x, below which rings start to play a significant role. In the studied ρ-window we find $0.125 \lesssim T_x \lesssim 0.14$, a value within the range of temperatures previously estimated in quasi two-dimensional simulations [171]. In the model studied in Chapter 2 we have defined in a similar way an onset temperature at which rings of a size are more abundant than chains of the same size, finding the value $T_{\text{onset}} \approx 0.75$. Rescaled by the minimum energy configuration of DHSs particles, this value would correspond to a DHS temperature ≈ 0.15. Fig. 3.14 shows also the fraction of rings as a function of temperature for a wide range of densities. As expected, there is a qualitative resemblance with Fig. 2.7: the curves are non-monotonic and the peaks grow upon lowering T and move to lower and lower densities.

We start by calculating the radius of gyration $R_g(s)$ of rings and chains of size s (Fig. 3.15). In the scaling limit, i.e. for large s,

$$R_g(s) = b \cdot s^{\nu}, \tag{3.10}$$

where the pre-factor b is model (and, in principle, T) dependent, while the exponent $\nu = 0.5$ for a random walk and $\nu = 0.588$ in the case of a self-avoiding random walk (SAW) [191]. $R_g(s)$ appears to be rather insensitive to T. Indeed, both the chain and the ring radii of gyration do not display any significant change upon lowering T. In the limit of large s, i.e. $s > 20$, for both chains and rings a power-law dependence (Eq. 3.10) sets in, but with an exponent $\nu \approx 0.7$

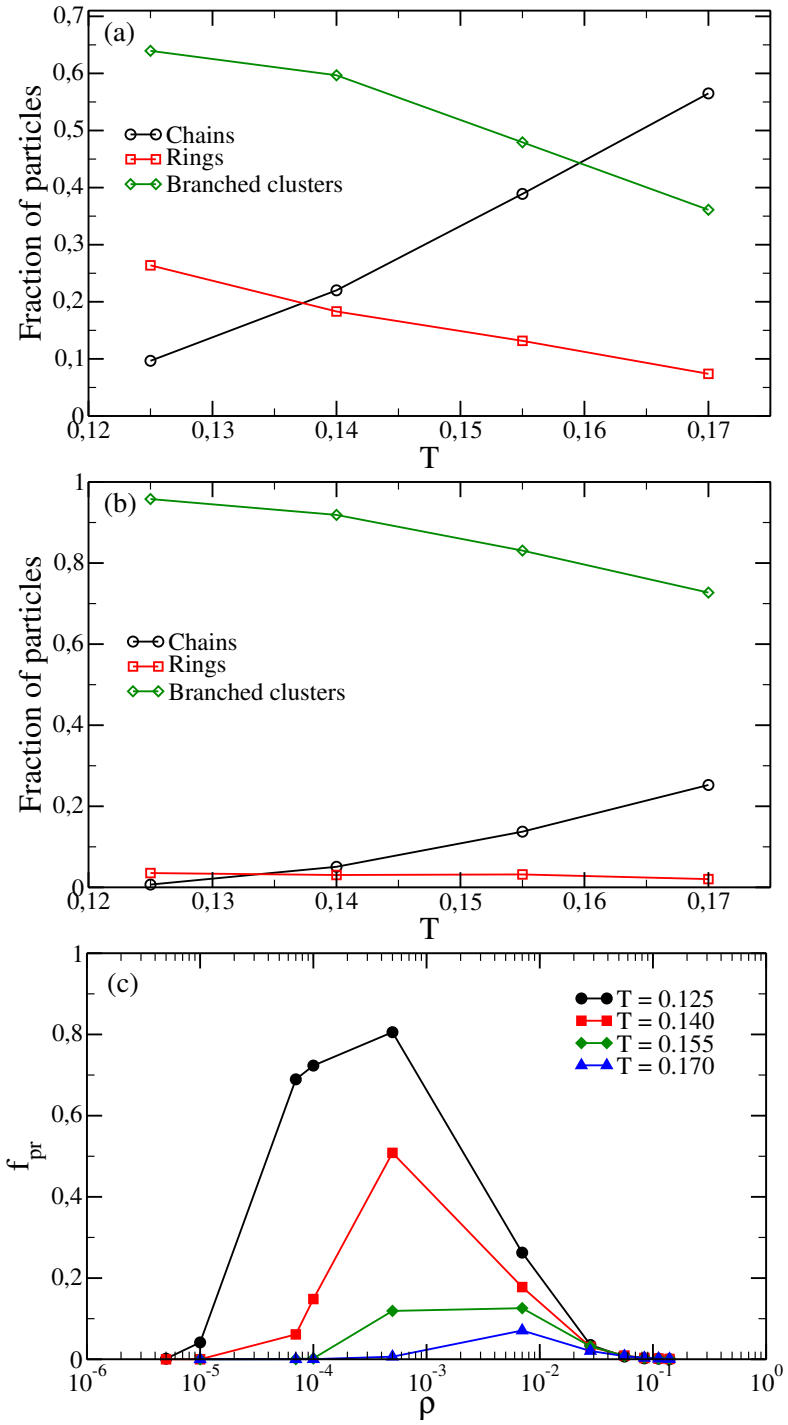

Figure 3.14: Fraction of particles in chains (circles), rings (squares) and branched structures (diamonds) for all the studied temperatures and (a) $\rho = 0.007$ and (b) $\rho = 0.028$. (c) Fraction of particles in rings for all the investigated T plotted against density.

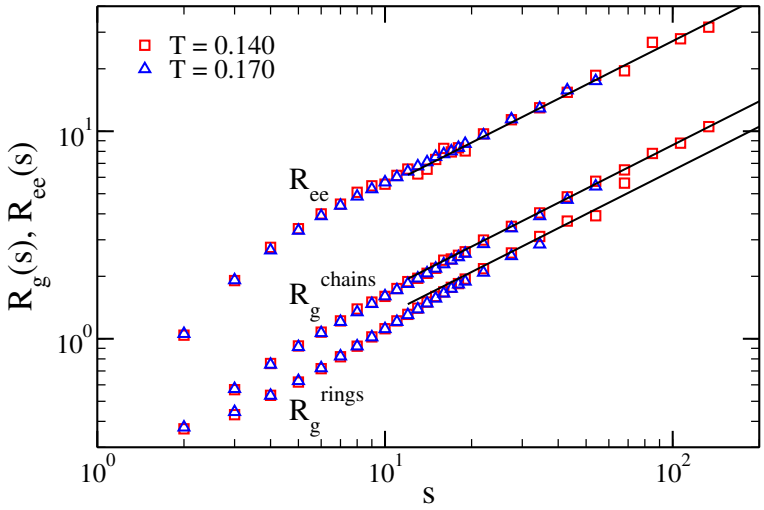

Figure 3.15: Radius of gyration $R_g(s)$ of rings $R_g^{\text{rings}}(s)$ and chains $R_g^{\text{chains}}(s)$ and mean end-to-end distance $R_{ee}(s)$ of chains at $\rho = 0.007$ for $T = 0.140$ and $T = 0.170$. Lines through the $T = 0.140$ data are power laws with exponent 0.7.

significantly larger than the one expected for a SAW chain. While we can not exclude the possibility that we are observing an intermediate cross-over region, the difference with the SAW exponent is significantly large to question if dipolar chains and rings belong to a different universality class. Indeed, in a previous study on the Stockmayer fluid [190] (i.e. in which the hard-sphere potential is replaced by a Lennard-Jones one) a similar value of ν was observed. The average end-to-end distance of chains, $R_{ee}(s)$, is also shown in Fig. 3.15. Like $R_g(s)$, $R_{ee}(s)$ is rather T independent and scales, within numerical uncertainty, with the same exponent $\nu \approx 0.7$.

To further characterize the chain geometry, we evaluate the angular correlation $\langle \cos(\theta) \rangle (n)$, where $\cos(\theta) = \hat{\boldsymbol{\mu}}_i \cdot \hat{\boldsymbol{\mu}}_j$, i and j are two particles belonging to the same chain, n is the "chemical" distance between them (i.e. the number of bonds separating the two particles) and angular brackets indicate an average taken over all pairs of particles (in the same chain) which are separated by n. We compare its behavior with the expected functional form

$$\langle \cos(\theta) \rangle (n) = \exp\left(-\frac{n}{l_p}\right) \tag{3.11}$$

where l_p is the chain persistence length. Fig. 3.16(a) shows both $\langle \cos\theta \rangle (n)$ and the best fits to Eq. (3.11). Table 3.1 reports l_p for all studied T. Unlike b in Eq. 3.10, lowering T from $T = 0.170$ to $T = 0.125$ results in a $\sim 100\%$ increasing

T	$\bar{l}$	l_p	e_0
0.125	64.2	4.6	-1.98
0.140	37.9	4.44	-1.92
0.155	15.2	2.87	-1.86
0.170	7.25	2.35	-1.79

Table 3.1: Average chain length $\bar{l}$, persistence length l_p and asymptotic mean internal energy per particle in chains and rings e_0 at $\rho = 0.007$ for all the studied temperatures.

of l_p.

To conclude, we analyse the mean squared magnetic moment of chains of size n, $\langle m_n \rangle$. Mendelev and Ivanov have provided a parameter-free expression for $\langle m_n \rangle$ in the dilute limit [192]

$$\langle m_n \rangle = \sqrt{n + 2\frac{K}{(1-K)^2}(n - 1 + K^n - nK)} \tag{3.12}$$

where $K = \coth(\beta/2) - 2/\beta$. Fig. 3.16(b) shows simulations results and predictions from Eq. 3.12 for $\langle m_n \rangle$ for two different T and $\rho = 0.007$. The theoretical predictions are in rather good agreement with numerical results.

Next we focus on the chain distribution $n_c(s)$, the number of chains of size s. Fig. 3.17(a) shows the first moment of the distribution, the average chain length $\bar{l}$ for all studied state points. Interestingly, at high T, $\bar{l}$ shows a non-monotonic dependence as a function of ρ. This is in stark contrast to what expected from an equilibrium polymerization process, in which only chaining is present. In mean field, $\bar{l}$ is predicted to scale as $\bar{l} \sim \rho^{1/2}$ [193]. The presence of a maximum stems from the equilibrium transformation of chains into branched structures. Indeed, within a mean-field description, the density of chains (ρ_c) increases with the density of chain ends (ρ_e) and decreases proportionally to the density of junctions (ρ_j), thus $\rho_c = k_1\rho_e - k_2\rho_j$, where k_1 and k_2 are proportionality factors and $\bar{l} \sim \rho_c^{-1}$. The corresponding scaling of ends and junctions is [55]: $\rho_e \sim \rho^{1/2}e^{-\epsilon_e/T}$, and $\rho_j \sim \rho^{3/2}e^{-\epsilon_j/T}$, where ϵ_e and ϵ_j is the energy cost of ends and junction respectively. By taking the derivative of the expression of the density of chains with respect to ρ, a minimum in ρ_c appears, resulting in a maximum for $\bar{l}$ at $\rho^* \sim e^{-(\epsilon_e - \epsilon_j)/T}$. Note that $\epsilon_e - \epsilon_j$ is the energy released when a chain end bonds to form a junction, suggesting indeed that the density maximum is related to the assembly of the chains in branched networks.

Fig. 3.17(b) shows the same $\bar{l}$ values but as a function of the reduced number density of particles in chains ρ_c. In this representation, the internal equilibrium

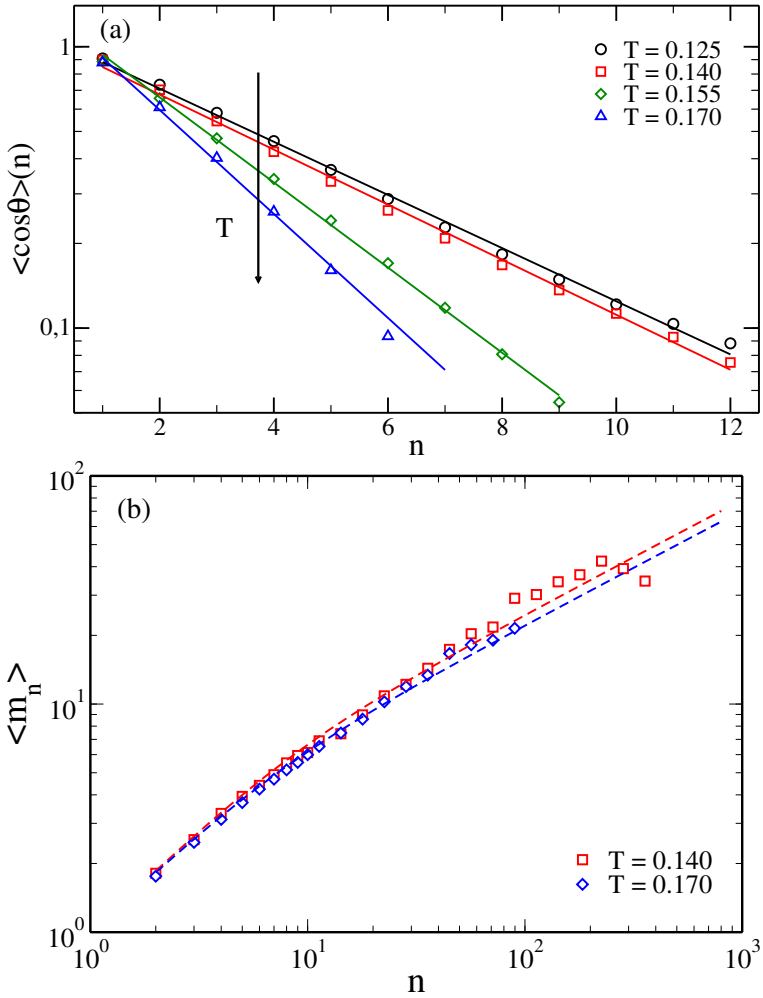

Figure 3.16: (a) Average angular correlation $\langle \cos\theta \rangle(n)$ between particles as a function of the chemical distance n along the chain (symbols). Lines are exponential fits to Eq. (3.11) restricted to small n. (b) Eq. (3.12) (dashed lines) and simulation results (symbols) for the mean squared magnetic moment $\langle m_n \rangle$ as a function of the chain length n for two different T.

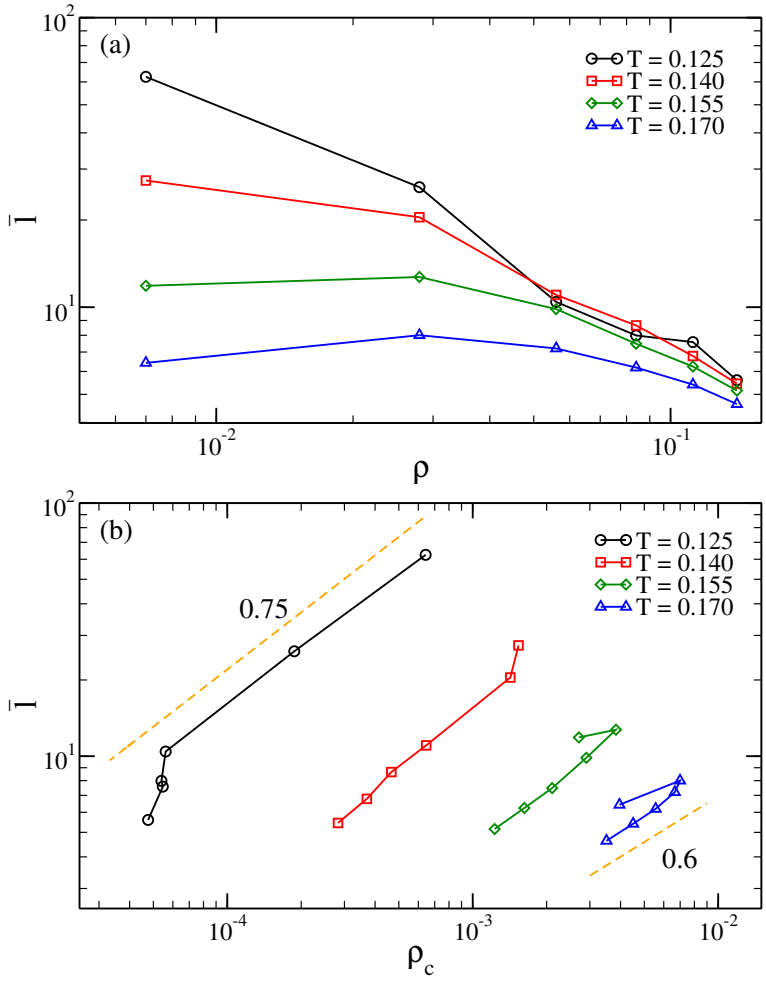

Figure 3.17: (a) Log-log plot of the average chain length $\bar{l}$ as a function of ρ for all the studied temperatures. (b) Same data but plotted as a function of the rescaled density N_c/V, where N_c is the number of particles in chains.

within the chain sub-system is considered and $\bar{l}$ returns to be a growing function of ρ_c, even if only in a finite interval. In this window, the ρ_c dependence of $\bar{l}$ is consistent with a power law behavior, but with a T-dependent exponent, between 0.6 and 0.75. As recalled before, in linear polymerization mean-field approaches predict in the low-T, low-ρ region an exponent 0.5 [133, 193, 194]. More accurate approaches predict, for dilute systems, 0.46 ± 0.01 [193].

The ring cluster size distribution $n_r(s)$ is shown in Fig. 3.18(a). The number of rings grows significantly on cooling, favoured by the additional energetic stability of rings as compared to chains of equal length provided by the additional bond. It has been suggested, on the basis of a Flory-Huggins-type mean-field

approximation [126] that $n_r(s)$ can be represented as

$$n_r(s) = c \exp\left(-\frac{s}{\bar{l}}\right) s^{-h} \tag{3.13}$$

where c is a normalization constant and h is a characteristic, T independent exponent. As shown in Fig. 3.18(a), imposing the known $\bar{l}$ value, it is possible to properly model $n_r(s)$ (for $s \gtrsim 15$) following Eq. (3.13).

By definition, the ratio between $n_c(s)$ and $n_r(s)$ is equal to the ratio of the respective partition functions. Three elements concur in controlling such ratio:

1. the different free energy contribution which arises from the different number of bonds in chains and rings of equal size. Assuming that at low T the dominant contribution is energetic, we expect that the ratio will depend on $e^{-\beta\Delta E(s)}$, where $\Delta E(s) \equiv s(e_r(s) - e_c(s))$ is the energy difference between a ring and a chain of equal size s;

2. the number of distinct modes which convert a ring to a chain, equal to the total number of bonds in the ring, s;

3. the different entropy of a chain compared to a ring, which can be approximated by the volume explored by the chain ends, $R^3_{ee}(s)$.

Fig. 3.18(b) shows the ratio,

$$r(s) \equiv \frac{n_r(s)}{n_c(s)} s R^3_{ee}(s) e^{-\beta\Delta E(s)} \tag{3.14}$$

for all the studied T. Apart from deviation at very small and very large s values, associated to the significant numerical noise at these extreme conditions, all the curves collapse on a single master curve for small and intermediate s. Even if the large s behaviour is plagued by numerical uncertainties arising from the small number of aggregates with such a large size, it appears plausible that $r(s)$ has a weak (or even no) dependence on s for large s, in agreement with previous results derived for equilibrium polymerization [126]. We also note that Eq. (3.14) is consistent with Eq. (3.13), since it predicts that the ring distribution function is the product of the chain distribution function (and hence of an exponential function decaying with $\bar{l}$) times a power law in s, which account for the asymptotic behavior of the end-to-end cube distance, the linear s dependence as well as any residual s dependence in the bond energy, which have not yet reached its asymptotic value (see Fig. 3.9).

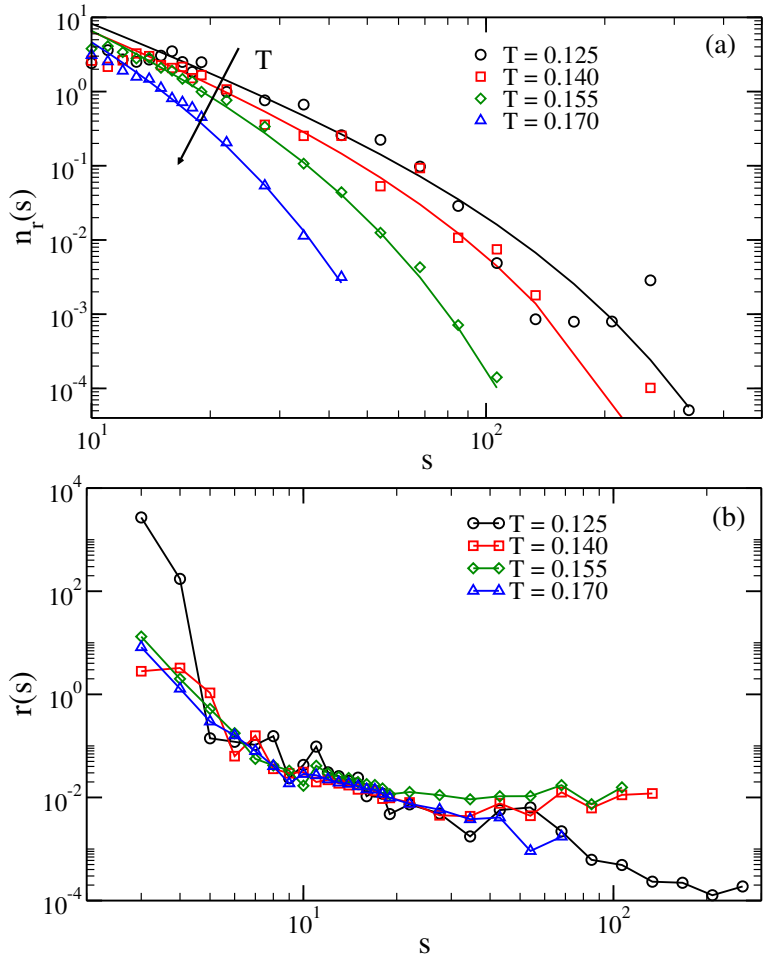

Figure 3.18: (a) Number of rings $n_r(s)$ of size $s \geq 10$ at $\rho = 0.007$ for every studied temperature (symbols). Lines are best fits to Eq. (3.13) limited to the region $s > 12$, with $\bar{l}$ taken from Table 3.1. Imposing h to be equal for all T results in a value of $h \simeq 2$, leaving only the amplitude c as a free, T dependent parameter. (b) $r(s)$ (see Eq. (3.14) for definition) for all studied T at $\rho = 0.007$. $\Delta E(s)$ is computed via the calculated $e_r(s)$ and $e_c(s)$ values for $s < 10$ and via the associated fit functions (see Section 3.3.4) otherwise.

3.4 Conclusions

In conclusion, in this Chapter we report an in-depth characterization of the thermodynamic and structural properties of DHS for $0.125 \leq T \leq 0.17$ and $0.007 \leq \rho \leq 0.14$, a window of T and ρ which, despite its relevance, was never explored before in simulation studies. To investigate such a region we have introduced specialized Monte Carlo biased moves and *ad hoc* simulation techniques, allowing us to effectively sample low-T and low-ρ configurations in equilibrium.

Our results show that the DHS fluid does not undergo any phase separation which was long thought to be hidden in this region. We also provide several hints on why it is so.

Indeed, the theoretical approaches developed to predict the low T DHS behaviour focused on the competition between bonding and chaining as the basic elements which control the thermodynamics of the system [55]. A new element brought in by the present study is the presence, in addition to chains and branched structures, of ring-shaped clusters. We have found that, at low T and ρ, rings become more probable than chains, since the additional energy gain of forming the additional bond which converts a chain into a ring compensates for the entropic loss associated to exploration of the volume available to the chain ends.

The dominance of rings over chains could affect the gas-liquid phase separation. Rings in fact have a net magnetic moment approximately null and are weakly interacting objects. More importantly, they deprive the fluid phase of chain ends, which were predicted to sustain the topological phase separation [55]. In general, an energy-driven phase-transition relies on the energy gain of the fluid to form aggregates to overcome the entropy penalty of condensation. With rings being the majority cluster at low-ρ, there is virtually no energy gain in the transformation of the gas of rings into a branched liquid-like network. This can be seen also by considering the surface tension of the aggregates, which, at low-T (where entropy plays a minor role), is proportional to the energy difference between particles on the surface of the aggregate and particles inside the aggregate. The loss of chain ends due to ring formation, lowers the surface tension of the fluid phase to value close to zero, potentially suppressing the driving force to condensation.

Of course, a phase transition is still an open possibility for T lower than the ones we have investigated. In this case the results of our study can be used to

guide future theoretical modelling in the search of a mechanism which could sustain criticality. If the phase transition between a liquid and a gas indeed exists at lower temperatures than the one we have been able to explore, the gas phase should be modelled as a phase rich in rings, instead of as phase rich in chain ends as it has been assumed until now.

As an example of such modelling, the mapping of the DHS properties into a patchy particle model appears to be a promising route, especially if the model can be extended to incorporate the chain-ring equilibrium as well as the branching at higher densities. In principle, it is possible to envision an asymmetric patchy particle model [143] in which the bond angle is comparable to the one characteristic of DHS (i.e. which give rises to similar $R_{ee}(s)$ and $\langle\cos(\theta)\rangle(n)$ as the one reported in Fig. 3.15 and 3.16) and in which the relative weight of rings and chains is controlled by the same ratio as the one discussed in Eq. (3.14). Such a model, if properly benchmarked against the data reported in this Thesis, could perhaps shed some light on the phase behaviour of very low T DHS systems. As a starting point, we have used the results gathered in this Chapter to tune the interaction parameters of the ring-forming, bivalent patchy particle model employed in Chapter 2.

Chapter 4

Dynamics of tetravalent patchy particles in the optimal network density

4.1 Introduction

The development of new techniques for particle synthesis makes it possible to produce surface-patterned colloids in bulk quantities [26–28, 30]. Theoretical and numerical results can help understanding how tuning the properties of these building blocks affects the resulting material. Indeed, patchy colloidal particles [8, 195, 196] continue to be the subject of an intense investigation, both experimentally and theoretically. There is indeed much expectation in the outcome of recent efforts of creating colloidal particles that interact via anisotropic potentials. The challenge faced by physicists, chemical engineers and material scientists is to organize these new geometries into structures for functional materials and devices via self-assembly, the spontaneous organization of matter into desired arrangements. The aim is to achieve – via the rational design of elementary building blocks – pre-defined specific, ordered or disordered, structures, shifting from the top-down to the bottom-up approach, in which effort is made in the direction of controlling particle shape and patterning.

Theoretical and numerical studies of the phase behaviour of patchy colloidal particles have been very much rewarding. Quite unexpectedly, a very rich framework for interpreting phenomena like thermoreversible gelation, the

competition between gelation and glass transition or the competition between condensation and polymerization [133] has been unravelled. New concepts like empty liquids [3, 39, 40], equilibrium gels [197, 198], unconventional gas-liquid phase diagrams [56, 182, 183] have been introduced and have been very fruitful in promoting further developments [63, 199]. One of the unexpected connections concerns the analogy between gelation in patchy colloids and glass formation in atomic and molecular network forming systems [63]. Indeed, studying the role of the valence M it has been disclosed that the packing fraction ϕ of the liquid coexisting with the gas decreases on decreasing M [39]. For the case of hard-sphere colloids with four patches, ϕ is of the order of 30%. This implies that, for larger values of ϕ, gas-liquid phase separation is not encountered on cooling. On progressively decreasing temperature, the average lifetime of a patch-patch bond increases and particles become arrested by being part of a long-lived network of bonds. Energetic bonds thus determine the slowing down of the dynamics and the approach to a non-ergodic state, a dynamic arrest that we call gelation [3, 197, 200, 201]. Dynamic arrest can thus be expected to be different from the one characterizing glassy states, where caging is controlled by excluded volume interactions. Limiting the valence is crucial for observing gelation, since it makes it possible to access, in equilibrium, the region of intermediate densities where packing does not play a major role. For spherically interacting colloids, the gas-liquid phase separation is much wider and the coexisting liquid density is found at $\rho\sigma^3 \approx 0.6 - 0.7$. Hence, if crystallization is pre-empted, a liquid of spherically interacting particles can be brought to low T to form a glass without phase separating only at very large densities.

In this Chapter we investigate in details the evolution on cooling of the self and collective dynamics of a model for tetrahedral patchy colloids in the gel region at a fixed value of the density, in the so-called optimal network density region. In this window of densities the system is expected to be able to form an ideal (fully connected) random tetrahedral network [64]. At lower densities, gas-liquid phase separation takes place, while at larger densities packing prevents the possibility of geometrically arranging all molecules with proper angular and distance constraints required to form bonds. Despite the different nature of the dynamical arrest process, driven by bonding and not by packing, the decay of the correlation functions for this four coordinated model resembles the one observed in glasses. In particular, the Arrenhius nature of the dynamics and the four-fold symmetry enforced by the patchiness of the surface makes this model qualitatively similar to other strong glass-forming

liquids such as silica and water [202, 203]. This study is aimed at providing information on the dynamics of a model tetravalent system in order to help design and produce experimental realisations of such limited-valence particles like the DNA tetramers investigated in Chapter 5.

4.2 Methods

We study a simple continuous model for tetrahedral patchy particles by means of Brownian dynamics simulations. A particle is modelled as a rigid body defined by the position of its center of mass and by $M = 4$ vectors indicating the locations of the four patches [3]. The interaction potential between particles 1 and 2 is

$$V(1,2) = V_{CM} + V_P \tag{4.1}$$

where V_{CM} is the potential acting between the centres of mass of the two particles, and V_P is the interaction between patches:

$$V_{CM}(1,2) = \left(\frac{\sigma}{r_{12}}\right)^m \tag{4.2}$$

$$V_P(1,2) = -\sum_{i=1}^{M}\sum_{j=1}^{M} \epsilon \exp\left[-\frac{1}{2}\left(\frac{r_{12}^{ij}}{\sigma\alpha}\right)^n\right]. \tag{4.3}$$

where σ is the particle diameter. The large value $m = 200$ is chosen to approximate the hard-sphere behaviour, the quantity $n = 10$ makes the exponential function resemble a square well, $\alpha = 0.12$ guaranties that the single bond per patch condition is satisfied and $\epsilon = 1.001$ fixes the absolute minimum at unitary depth. The distance between the centres of particles 1 and 2 is indicated as r_{12} while the distance between patches on different particles with the symbol r_{12}^{ij}. Bond forces thus act on surface spots allowing momenta which can induce particle rotations.

The parameters entering in Eqs. (4.2) and (4.3) have been chosen in such a way that the resulting potential has a depth $u_0 = -1$ and it resembles an hard sphere plus square well potential, allowing greater flexibility in the study of the dynamics of these systems compared to step-wise potentials. Note that, while in the Kern-Frenkel potential [144] the interaction range and the angular width of the bond can be independently controlled, in the present continuous potential the patch-patch interaction depends only on the patch-patch distance

and hence the patch-patch interaction range and angular width are coupled.

We perform Brownian dynamics simulations in the NVT ensemble using 10000 particles in a cubic box of size $L = 26$ with periodic boundary conditions. In the following, the energy unit is chosen to be the depth u_0 of the potential, the length unit is chosen to be the colloids diameter σ and time is in units of $\sigma\sqrt{m/u_0}$, where m is the mass of the colloids. The Brownian algorithm used in the simulations is described in the appendix of Ref. [3]. Here we summarize its features. A Velocity-Verlet integrator with an integration time $\delta t = 0.001$ is used to integrate the equations of motion. To model Brownian diffusion, we define a probability p for each particle to undergo a random collision every N time steps. By tuning p it is possible to obtain the desired free particle diffusion coefficient D_0 using the relation

$$D_0 = \frac{k_B T N \delta t}{m}\left(\frac{1}{p} - \frac{1}{2}\right) \tag{4.4}$$

In the simulation units the chosen translational bare diffusion coefficient is $D_0^T = 0.01$ and the corresponding rotational diffusion coefficient id $D_0^R = 0.03$ (so that $D_0^R/D_0^T = 3$, as expected for non-slip particles). These values fix p^T and p^R for each temperature.

The average time between two random collisions is given by

$$\Delta t = \frac{N\delta t}{p} = \frac{k_B T N \delta t + 2mD_0}{2k_B T} \tag{4.5}$$

thus our simulations follow a Newtonian dynamics for $t < \Delta t$ and a Brownian dynamics after that time.

In order to equilibrate at the lowest temperatures we use a version of the code that runs on GPUs using CUDA [204], as reported in Chapter 6. The simulations were performed on Tesla C2050 GPUs.

As discussed in Chapter 6, the performances achieved on GPUs depend heavily on interaction details (such as cut-off distances), number of particles and density. For the state points investigated in this work ($N = 10000$, $\rho\sigma^3 = 0.57$) we obtain a 30x speed-up with respect to a Xeon E5620 (single core).

We have used up to 10^{10} MD steps for equilibration and from 10^7 to $2 \cdot 10^9$ MD steps for data generation, depending on the temperature.

In Brownian dynamics simulations particles are subject to a random force which accounts for the collisions between colloidal particles and the solvent. Even if the random force has a zero mean, the position of the center of mass

of the system becomes a random variable, again with zero mean. In standard simulations the random motion of the center of mass (COM) is negligible and the particle dynamics is weakly affected by this random process. In the study of very long simulations, as the ones reported here (which extend to 10^9 integration time steps), the motion of the center of mass can be quite substantial and can produce artifacts in the evaluation of dynamical quantities. For this reason, in all data presented in this work the trajectories of the single particles have been corrected to subtract the center of mass motion. Care need to be taken in the analysis of Brownian dynamics trajectories in glassy states, especially now that the increased power of GPUs for scientific application is opening the possibility of investigating glassy states via lengthy simulations.

4.3 Results

4.3.1 Static

To properly frame the investigation of the dynamics, we start by showing in Fig. 4.1(a) the potential energy per particle as a function of T. The energy has the typical sigmoidal shape characteristic of bond interactions, reminiscent of the two-state behaviour of the bonds (broken or formed). On cooling, the system changes from a collection of isolated clusters to a percolating network to an essentially fully bonded configuration, with a few isolated monomers, detaching from the infinite cluster, as indicated by the cluster size distribution, reported in Fig. 4.1(b). Two particles are considered connected (and hence belonging to the same cluster) if their pair interaction energy is lower than -0.5. Below $T = 0.15$, a large fraction of particles belongs to the infinite cluster and at the lowest investigated T, more than 99% of the particles are in the infinite cluster (see inset of Fig. 4.1(b)). Hence, at low T the system can be visualized as a percolating network which incorporates most of the particles.

Fig. 4.2 shows the static structure factor for the lowest investigated temperatures. Beside the main peak at $q\sigma \approx 8$, the structure factor shows a pre-peak characteristic of tetrahedral networks, where the slowest collective modes are found. In addition, we note that the T dependence of $S(q)$ tends to saturate at low T ($S(q)$ at $T = 0.10$ and $T = 0.105$ are identical within the noise) suggesting that the system structure does not significantly evolve any longer with T. This saturation of $S(q)$ at small T has been interpreted as evidence of an equilibrium gel state [39, 205].

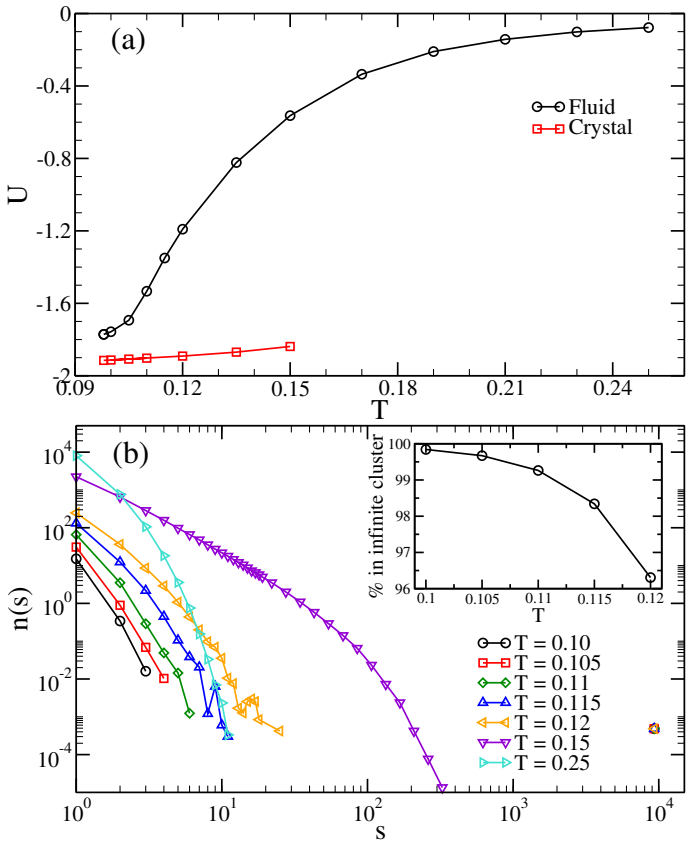

Figure 4.1: (a) Potential energy per particle U as a function of temperature (black circles). Also shown for comparison is the potential energy per particle in a diamond crystal structure (red squares). (b) Number of clusters $n(s)$ of size s for different temperatures for a system of 10000 monomers. For $T < 0.15$ the system always contains a percolating cluster (disconnected points at $s \approx 10^4$). Inset: percentage of particles that are in the infinite cluster (P_∞) as a function of T.

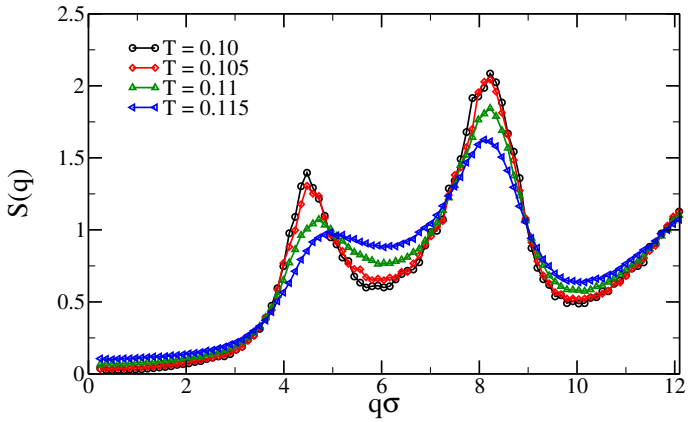

Figure 4.2: Structure factors for the lowest temperatures.

4.3.2 Bond lifetime

The bond-bond autocorrelation function $C_b(t)$, defined as the probability that a bond existing at $t = 0$ exists also at time t, provides a quantification of the typical microscopic time, setting the scale for the dynamics, separating the (short) time scale in which the dynamics takes place at fixed bonding pattern from the (long) time scale where dynamics is intrinsically connected to bond breaking events. Fig. 4.3(a) shows $C_b(t)$ for all the investigated T, showing that more than five order of magnitudes in bond lifetime are properly explored. The decay of the correlation function can be fitted with a stretched exponential function, $e^{-(t/\tau_b)^{\beta_b}}$, and the values of β_b are reported in the inset, while the T dependence of τ_b is shown in Fig 4.3(b). The decay is clearly stretched, suggesting that the different local bonding environments have a role in the process of bond breaking. From the fit, an average bond time can be calculated as $\langle \tau_b \rangle = \frac{\tau_b}{\beta_b} \Gamma\left(\frac{1}{\beta_b}\right)$, where $\Gamma(x)$ is the Gamma function.

4.3.3 MSD and D

Fig. 4.4(a) shows the mean square displacement (MSD) for all the investigated T. Above and around percolation, no plateau in the time dependence of the MSD is observed. Indeed, close to $T = 0.15$, the lifetime of the bonds is still very short and no dynamical signatures of the presence of a transient infinite cluster are observed. Below percolation, an inflection develops which gives rise to a plateau which increases on further cooling. The height of the inflection point significantly changes with T, signalling that bonding progressively reduces the cage volume. At the lowest T, the MSD becomes comparable to the square of the bonding distance, suggesting that in a (almost) fully bonded locally tetrahedral structure, bond confinement can be quite effective.

The diffusion coefficient D, evaluated from the long time limit of the mean square displacement (MSD $= 6Dt$) is shown in Fig. 4.4(b). D is clearly super-Arrhenius around the percolation temperature. In this T-interval indeed the structure of the system changes significantly, since particles first aggregate into larger and larger clusters and, beyond percolation, join more and more the spanning cluster. As discussed later on, upon entering well inside percolation the structure of the system reaches its equilibrium gel state and no further significant structural changes take place. At these low T, D shows an apparent Arrhenius behaviour, with an activation energy of about -4.5, a value slightly larger than the energy required to completely break four bonds. The Arrhenius

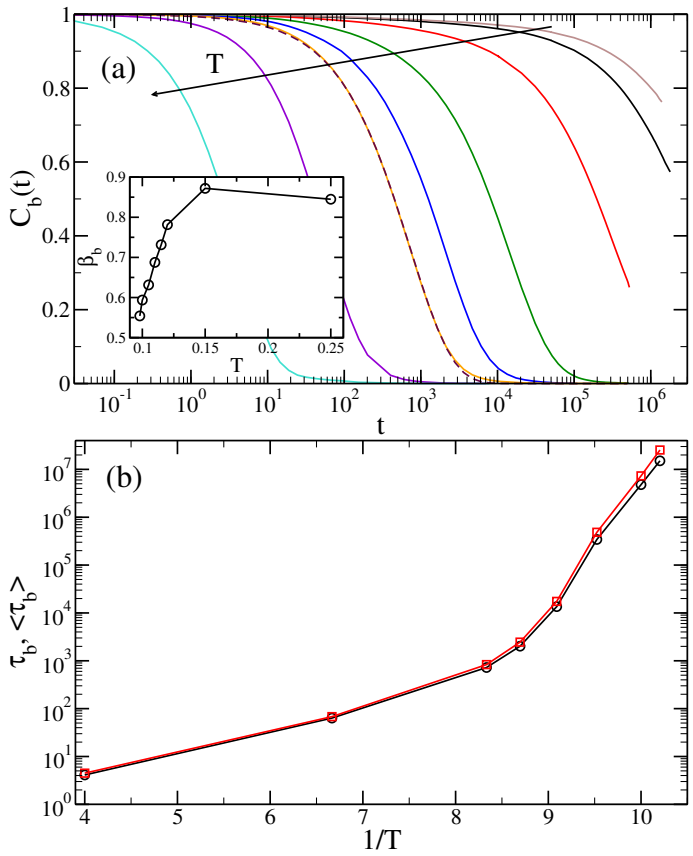

Figure 4.3: (a) Bond-bond autocorrelation function for different temperatures ($T = 0.098, 0.10, 0.105, 0.11, 0.115, 0.12, 0.15, 0.25$). A stretched exponential fit to $C_b(t)$ at $T = 0.12$ is also included (dashed line). Inset: values of the fit parameter β_b (see text) for different temperatures. (b) Values of the fit parameter τ_b (black circles) and the average bond time $\langle \tau_b \rangle$ (red squares, see text) for different temperatures.

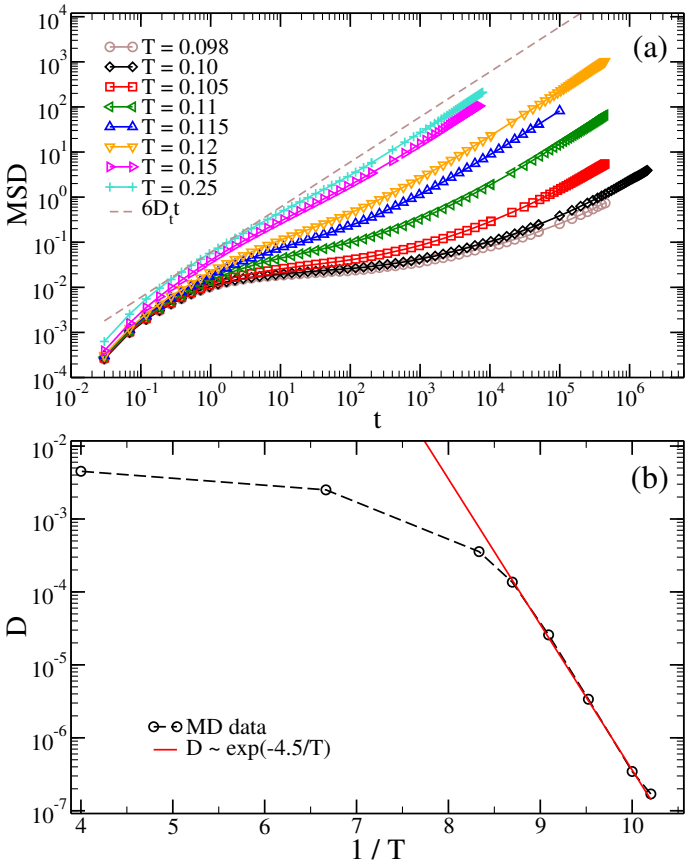

Figure 4.4: (a) Mean square displacement for different temperatures. The dashed line shows the expected time dependence of the diffusive behaviour at long times for monomers (for which $D_t = 0.01$). (b) Diffusion coefficient D extracted from the slope of the MSD at long time for different temperatures (black circles). The red line is an Arrhenius fit performed over the five lowest temperatures.

dependence classifies the present model in the category of strong glass forming systems [206], which includes all tetrahedral network fluids. In this respect, the present results confirm once more that there is a strong connection between the insurgence of an open local structure held together by strong directional forces and the observation of an Arrhenius dynamics. It is also interesting to observe that a similar value has been recently reported in the study of ST2 water at the optimal network density [207]. Values of the activation energy of the order of four bonds have also been observed in a model of tetrahedral DNA constructs [67] and in a primitive model for water [64], suggesting that the mechanism for microscopic dynamics in tetrahedral networks shares common features.

4.3.4 Self dynamics

To analyse the tagged particle motion in the wave vector $\vec{q}$ space, we evaluate the self-intermediate scattering functions $F_s(q,t)$, defined as

$$F_s(q,t) = \frac{1}{N}\sum_{i=1}^{N} < e^{-i\vec{q}\cdot(\vec{r}_i(t)-\vec{r}_i(0))} > \tag{4.6}$$

where $\vec{r}_i(t)$ is the position of the center of particle i at time t. The behaviour of the correlation functions is shown in Fig. 4.5 as a function of T for two different wave-vector values ((a) and (b) panels) and as a function of q at $T = 0.10$ ((c) panel). To help comparing the characteristic time-scales, the bond autocorrelation function is also reported. The long time decay of these functions can be rather well modelled via stretched exponentials,

$$F_s(q,t) = f_q^s e^{-(t/\tau_s(q))^{\beta_q^s}} \tag{4.7}$$

where f_q^s plays the role of the non-ergodicity factor, $\tau_s(q)$ is the characteristic decay time and β_q^s is the stretching exponent. The wave-vector dependence of f_q^s, $\tau_s(q)$ and β_q^s is reported in Figures 4.6 and 4.7. The non-ergodicity parameter shows the typical gaussian shape, but with an amplitude that is clearly T dependent, confirming that the cage volume decreases on cooling. A similar effect is also observed in glasses, but only below the so-called mode-coupling critical temperature [208], suggesting that somehow network liquids remain sufficiently fluid to be observed well below the point where dynamics crosses from power-law to Arrhenius. The stretching exponent is also T dependent and varies significantly on varying q. At small q, where dynamics has to convert to diffusive dynamics, β_q^s approaches one and $\tau_s(q) \sim q^{-2}$. Finally, Fig. 4.7(a) shows $\tau_s(q)$ and Fig. 4.7(b) shows $\tau_s(q) \cdot q^2$. It also shows the corresponding lifetime of the bond. Interestingly, the crossing between τ_b and $\tau_s(q)$ takes place to larger and larger q values on cooling, suggesting that the dynamics on smaller and smaller time-scales becomes more and more slaved to the bond breaking process. Only when time has become longer than the bond-breaking time particles are able to restructure themselves and relax the density fluctuations. Fig. 4.7(b) also shows that at the lowest investigated T, the approach to the diffusive limit is not clearly reached within the wave-vector range which can be explored by our simulation.

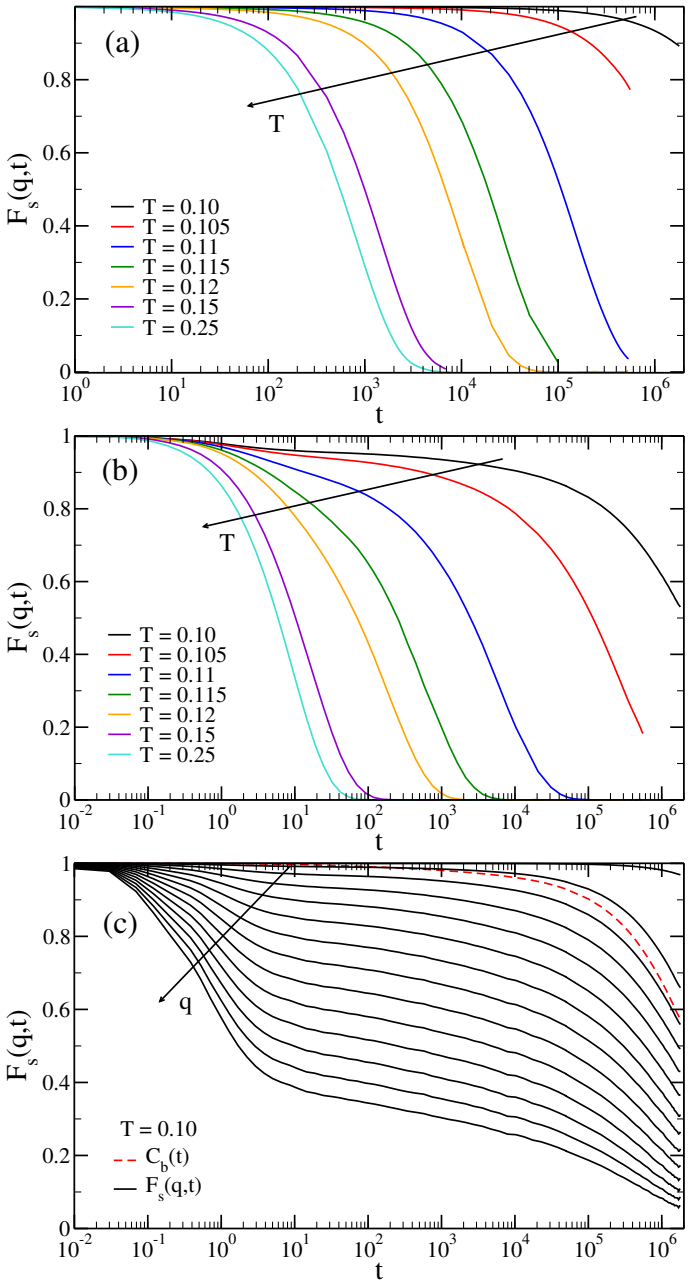

Figure 4.5: (a) $F_s(q,t)$ at different temperatures for $q\sigma = 0.48$. (b) $F_s(q,t)$ at different temperatures for $q\sigma = 4.59$ (corresponding to the first peak in the $S(q)$, see Fig. 4.2). (c) $F_s(q,t)$ at $T = 0.10$ for different values of the wave vector $q\sigma$ (0.24, 2.1, 3.9, 5.7, 7.5, 9.3, 11, 13, 15, 17, 18, 20, 22, 24). Also shown is the $C_b(t)$ at the same temperature (dashed red line).

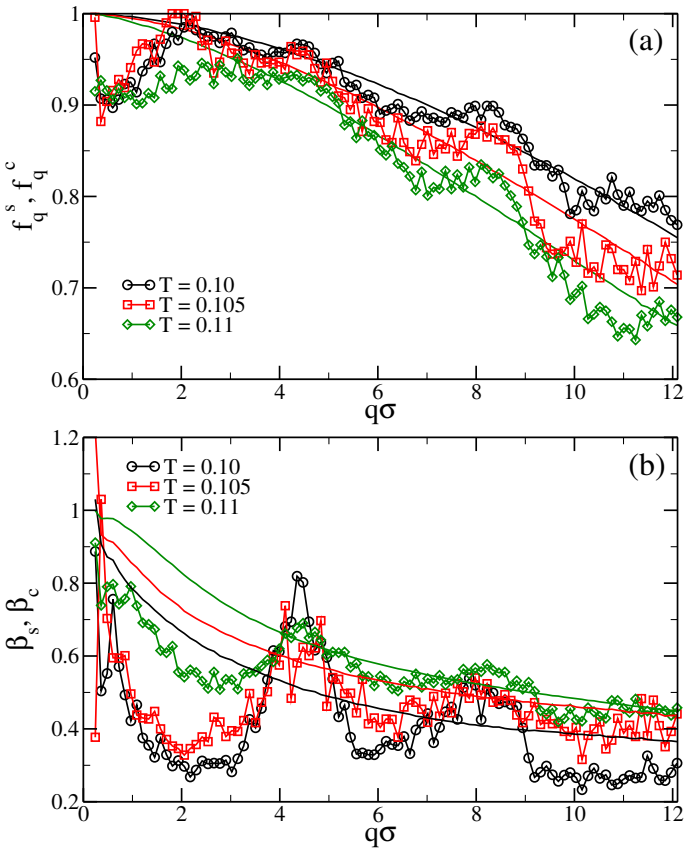

Figure 4.6: Fit results for the self (solid lines) and collective (symbols) intermediate scattering functions for different temperatures as functions of the wave vector. (a) Non ergodicity factors. (b) Stretched exponents.

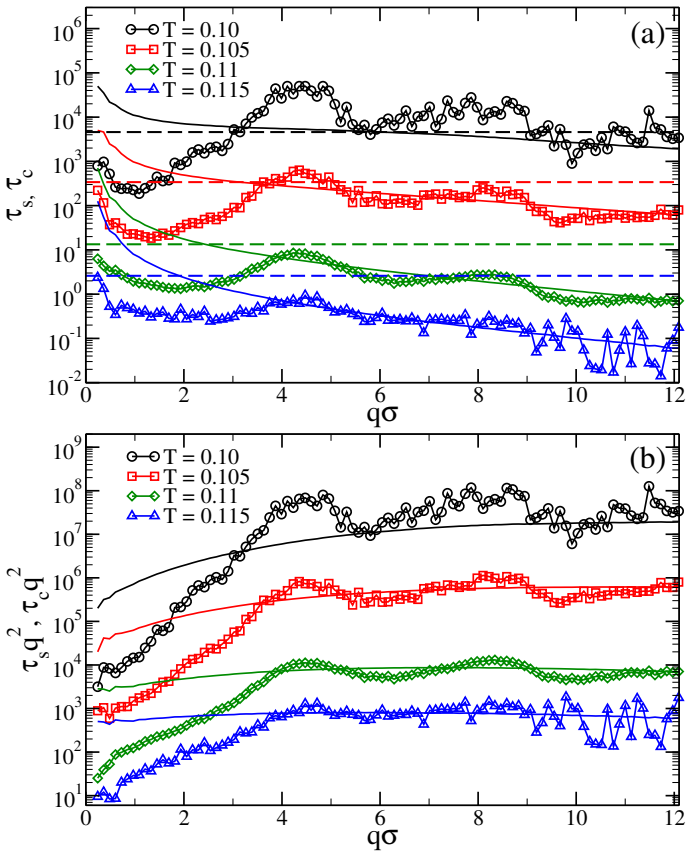

Figure 4.7: Fit results for the self (solid lines) and collective (symbols) intermediate scattering functions for different temperatures as functions of the wave vector. (a) Characteristic decay times (dashed lines are τ_b, added for comparison). (b) Characteristic decay times multiplied q^2.

4.3.5 Collective dynamics

To analyse the collective particle motion in wave vector $\vec{q}$ space, we evaluate the coherent intermediate scattering functions $F_c(q,t)$, defined as

$$F_c(q,t) = \frac{1}{N} < \sum_{i,j=1}^{N} e^{-i\vec{q}\cdot(\vec{r}_i(t)-\vec{r}_j(0))} > \tag{4.8}$$

As for the self-case, we show in Fig. 4.8 the q and T dependence of the collective correlation function. The data are significantly more noisy, reflecting the absence in the average over the distinct tagged particles. Despite the noise, some trends are clear and worth discussing. At low temperatures, very long times and small wave vectors, $F_c(q,t)$ shows an oscillatory behaviour which has been tentatively attributed to the presence of acoustic sound modes [197]. If we use Eq. (4.7) to fit these curves, we can note that also the collective non ergodicity parameter shows a clear T dependence at low T. Similarly to what has been observed for glasses [209–211], the values of f_q^c indeed oscillate around the self ones, in phase with the position of the peaks of the structure factor $S(q)$, shown in Fig. 4.2.

Interestingly, the decay of the density fluctuations does not always requires the breaking of bonds. For example, at $T = 0.115$, $\tau_c(q)$ is always smaller than τ_b, suggesting that the decay of the density fluctuations at small q happens at a fixed network structure. On further cooling, $\tau_c(q)$ and τ_b get closer, and at $T = 0.10$ $\tau_c(q) \gtrsim \tau_b$ for intermediate and large values of q.

This means that the breathing modes of the network, diffusive in nature, are of sufficient amplitude to relax the density fluctuations at large wave-length. Only at very low T the gel becomes so stiff that the decay of the density fluctuations takes place on a time scale comparable or longer than τ_b. This is more clearly shown in Fig. 4.9: at $T = 0.115$ the bond-bond autocorrelation decays faster than the density fluctuations at $q\sigma = 4.59$; at $T = 0.105$ the two times are similar while at $T = 0.10$ the opposite behaviour is observed and the decay of the density fluctuations requires the breaking of the network to take place.

4.4 Conclusions

In this chapter we have reported a study of the self and collective dynamics of a simple tetrahedral patchy model for a colloidal particle decorated by four attractive sites, located on the vertex of a tetrahedron. The shape and range of the site-site interaction is chosen in such a way that particles can form at

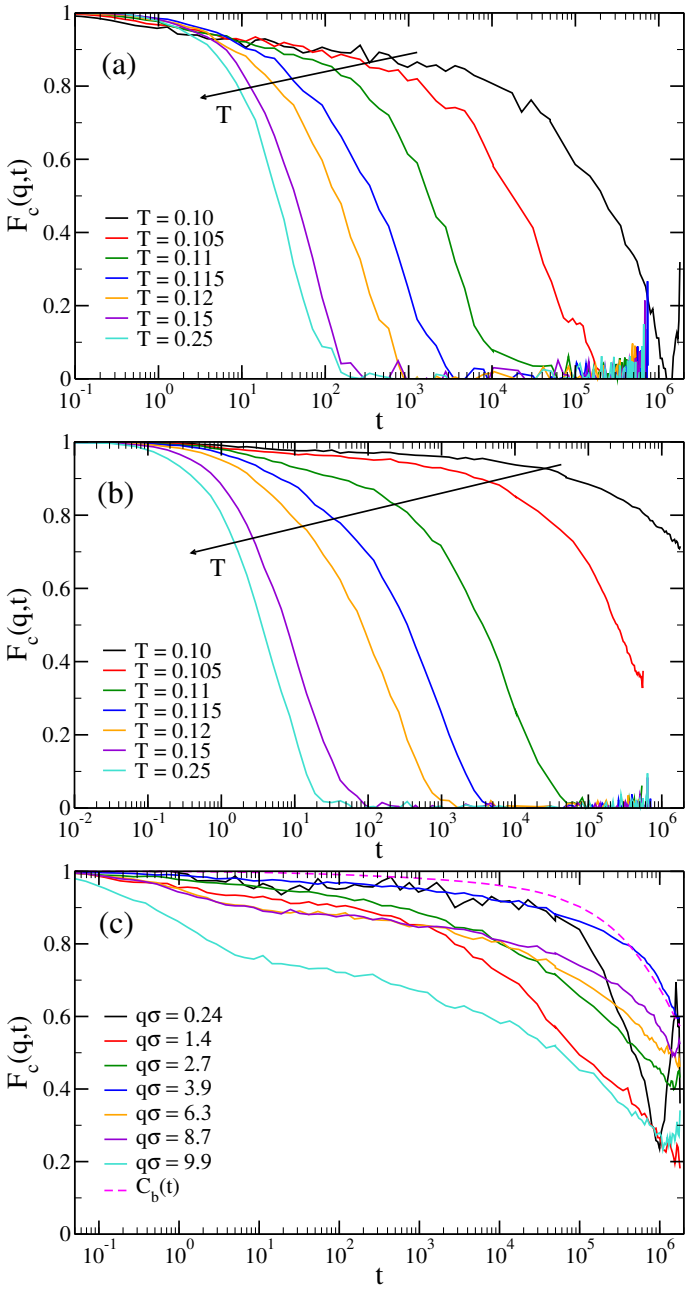

Figure 4.8: (a) $F_c(q,t)$ at different temperatures for $q\sigma = 0.48$. (b) $F_c(q,t)$ at different temperatures for $q\sigma = 4.59$. (c) $F_c(q,t)$ at $T = 0.10$ for different values of the wave vector. Also shown is the $C_b(t)$ at the same temperature (dashed magenta line).

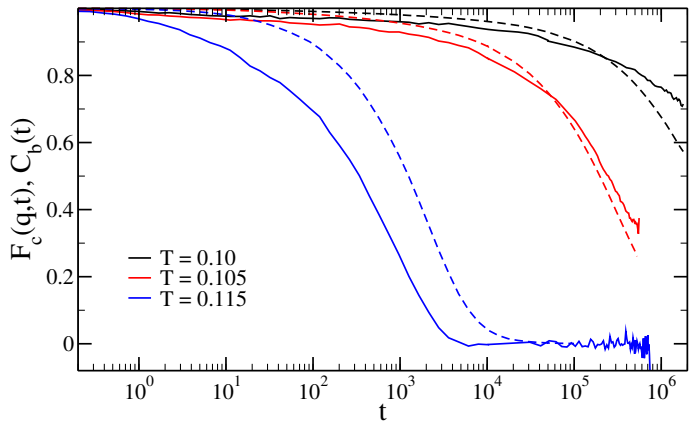

Figure 4.9: $F_c(q,t)$ for $q\sigma = 4.59$ (solid lines) and $C_b(t)$ (dashed lines) at three different temperatures.

most one bond per site. On cooling, the system progresses from a collection of isolated clusters to an essentially fully bonded tetrahedral network, in which most of the particles are engaged in four bonds. We have investigated the model at a fixed density, in the optimal network density region. Indeed, the presence of a limited number of strong directional interactions determines a limited range of densities which are compatible with the possibility of satisfying all possible interacting sites. This optimal density region is limited at low density by the presence of a gas-liquid coexistence and at high density by increasing packing, preventing the possibility of approaching the fully bonded network state. Along this isochore, the dynamics progressively slows down, first with a super-Arrhenius T dependence (around percolation), then crossing to an Arrhenius dependence at low T. The behaviour is similar to the one reported for silica and water, where also a cross-over from super-Arrhenius to Arrhenius has been observed in connection to the establishment of an extensively connected network [202, 203]. In the case of silica this crossover has been interpreted as a manifestation of the Mode-Coupling temperature [212]. In the present model, in which bonding is unambiguously defined, it appears that Arrhenius dynamics sets in when most of the particles belong to the spanning cluster. Interestingly, the comparison between the time-scales of bond-breaking events and diffusional processes clearly shows that diffusion over long distances (as detected by $F_s(q,t)$) is slaved to the bond lifetime and a truly diffusional process (such that $\tau_s q^2$ is approximately constant) can be observed at very small wave vectors only. The characteristic time-scale of the collective dynamics shows oscillations in phase with the structure factor, similarly to what has been found

in the case of atomic and molecular glass formers [209–211]. Interestingly, the decay of the density fluctuations does not always require the breaking of bonds. Only at very low T the gel has become so stiff that the decay of the density fluctuations, even on length scales comparable to the particle size, requires the preliminary breakdown of the bond network. Under these conditions, the self and collective dynamic become slaved to the time scale set by τ_b.

Chapter 5

DNA tetramers

5.1 Introduction

In the previous Chapters of this Thesis we have shown that, by carefully designing the inter-particle interactions, it is possible to produce materials with novel interesting and technologically relevant properties. Now we show that the obtained theoretical results can be extended to real systems by employing DNA as a building block.

Thanks to the highly specificity of the Watson-Crick pairing mechanism, DNA can be used to design and build new materials, whose basic components have tunable mutual interactions and predictable geometries. Indeed, recent developments in DNA synthesis and nanotechnology have made it possible to exploit DNA as a building block to produce 2D and 3D crystals [213, 214], complex structures [214, 215], hierarchical self-assembly of tiles [213] and self-assembly of strands into large structures, i.e. DNA origami [216]. DNA can also be used to functionalise colloids by grafting single-stranded DNA (ssDNA) molecules on the surface of the particles [67, 70, 217]. By choosing strands terminating with complementary sequences, the particles can bind to each other via DNA hybridisation and self-assemble into disordered or ordered structures [67,218–221]. Despite the considerable progress made in the synthesis of these particles, producing particles with a discrete number of ssDNA in large quantities remains a challenge [222–224].

In order to study an experimentally realisable valence-limited system we employ DNA sequences which undergo a two-step self-assembling process by forming tetravalent nano–sized DNA constructs, i.e. DNA tetramers, that,

Name	Sequence
Tetr1	CTACTATGGCGGGTGATAAAAACGGGAAGAGCATGCCCATCCACGATCG
Tetr2	GGATGGGCATGCTCTTCCCGAACTCAACTGCCTGGTGATACGACGATCG
Tetr3	CGTATCACCAGGCAGTTGAGAACATGCGAGGGTCCAATACCGACGATCG
Tetr4	CGGTATTGGACCCTCGCATGAATTTATCACCCGCCATAGTAGACGATCG

Table 5.1: Single-strand sequences designed to self-assemble into tetramers. Spacers are color-coded in grey and sticky ends in black.

upon lowering the temperature, bind together via sticky ends. In order to clearly separate the temperatures at which the two processes take place we use 20-base-sequences for the arms and palindromic 6-base-sequences for the sticky ends. These two parts are separated by nucleotides, acting as spacers, in order to increase flexibility. Table 5.1 contains the four ssDNA sequences. We use the same sequences employed by the group of prof. Bellini in Milan [225]. The hierarchical self-assembly process is shown in Fig. 5.1.

In this Chapter we use the realistic, coarse-grained DNA model developed by Ouldridge *et al.* [110,112], presented in Chapter 1, to simulate bulk systems of DNA tetramers at different concentrations and temperatures. In particular, we focus on the region of the phase diagram where the DNA constructs undergo a gas-liquid-like phase separation and we study the structure and the dynamics as the system approaches the spinodal curve.

5.2 Methods

The coarse-grained DNA model we employ has been already introduced in Section 1.2.1. Here we recall that the interaction forms and parameters are chosen to reproduce structural and thermodynamic properties of both single- and double- (dsDNA) stranded molecules of DNA in B-form. All interactions between nucleotides are pairwise and, in the last version of the model [110], continuous and differentiable.

The interactions between nucleotides account for excluded volume, backbone connectivity, Watson-Crick hydrogen bonding, stacking, cross-stacking and coaxial-stacking (see Fig. 1.1). The interaction parameters have been adjusted in order to be consistent with experimental data [110, 116, 226]. In addition, the model assumes conditions of high salt molarity (0.5 M). Lengths are expressed in units of $\sigma = 8.518$ Å and temperatures in degrees Celsius. In computing the concentration we use an average nucleotide mass of $m = 330$ *Da*.

In order to extract bulk properties from simulations, we investigate sys-

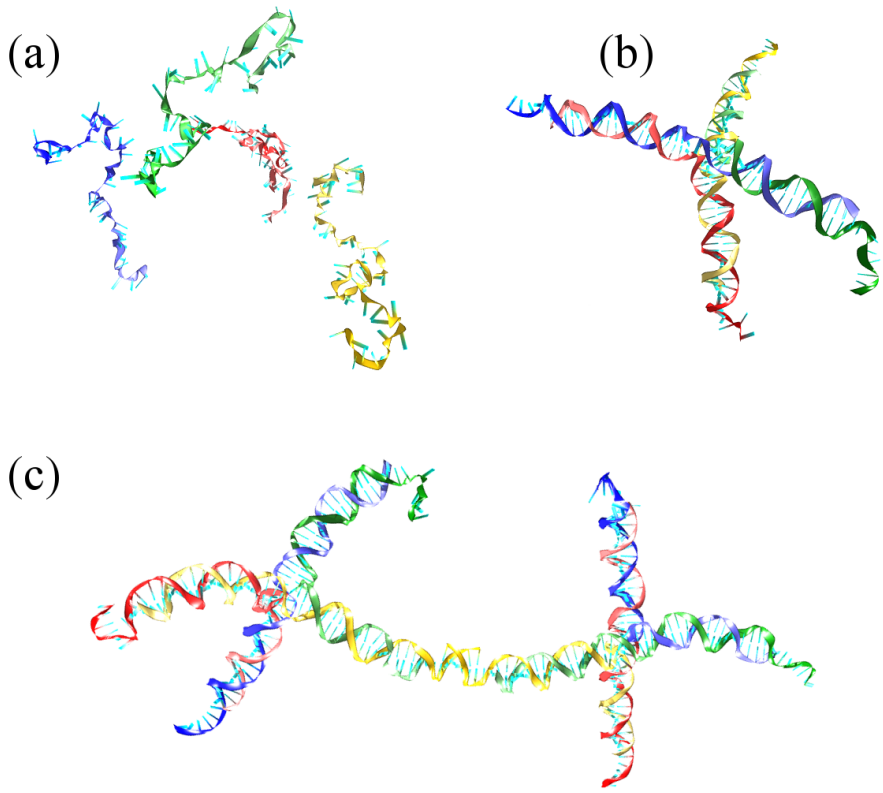

Figure 5.1: The investigated two-step self-assembly process. (a) At very high temperatures ($T \gtrsim 80\ C^\circ$) there are no hydrogen bonds between the strands. (b) For $80\ C^\circ \gtrsim T \gtrsim 50\ C^\circ$ single strands start to hybridise and tetramers are formed. (c) Upon further cooling, tetramers are linked together by hybridisation of the sticky ends. Each different strand color corresponds to a different sequence (see Table 5.1). All the sticky ends have the same palindromic sequence and hence they can bond to each other regardless of the strand they are part of.

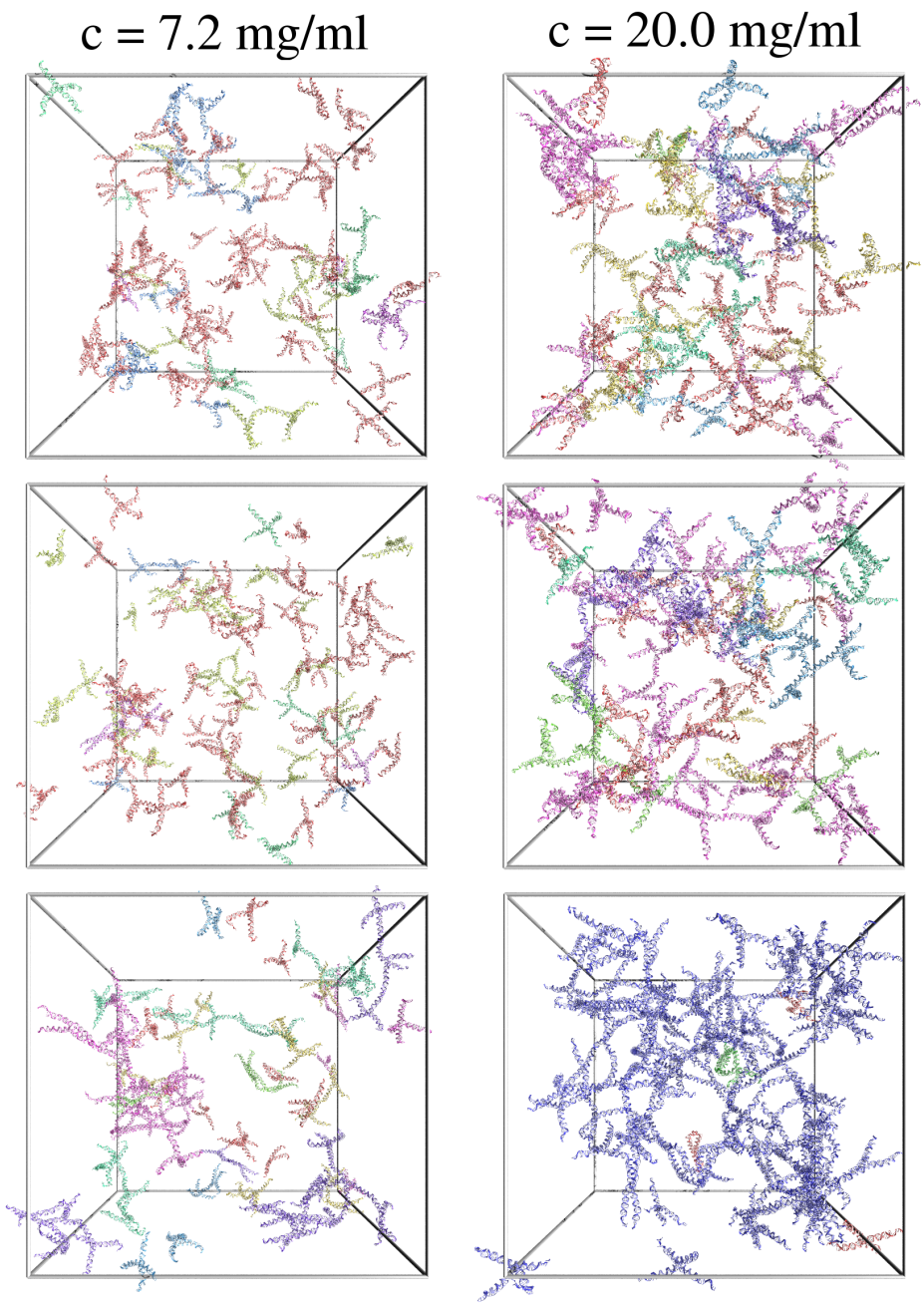

Figure 5.2: Snapshots taken from simulations at concentration $c = 7.2$ mg/ml (left panels) and $c = 20.0$ mg/ml (right panels) at different temperatures. Tetramers' colour depends on the size of the cluster they are part of. At high temperatures (top panels) there are mostly monomers (depicted in red), and the system is homogeneous. Upon lowering T, the system starts to form clusters but remains homogeneous (middle panels). At the lowest T (bottom panels) the system is inhomogeneous and, at high concentration, there is a percolating cluster that spans the whole simulation box.

tems made up of 100 DNA tetramers. Since each tetramer is composed of 196 nucleotides, we need to simulate 19600 particles interacting through a very complicated and numerically-intensive potential. By harvesting the computational horse-power of modern GPUs, as described in Chapter 6 and shown in Fig. 6.5, we are able to boost up performances by a factor $40 - 50$, if compared with CPU simulations. Equilibration and production simulations have been run up to 10^9 MD steps for each state point, corresponding to ~ 10 µs of real time. Taking into account the speed-up, each investigated system would have taken up a few years on a CPU single core.

We construct the initial tetramers by putting a high concentration ($c \approx 100$ mg/ml) of an equal number of constituent strands (see Table 5.1) in the simulation box. We then simulate the system at $T = 60\ ^\circ C$, since the melting temperature (i.e. the temperature at which the duplex yield is 0.5) of twenty-average-base strands is $\approx 80\ ^\circ C$ [110]. After the some time we have observed the formation of a few complete tetramers, which have been extracted to generate the initial configurations for the bulk simulations.

We simulate systems in a temperature range $39\ C^\circ \leq T \leq 48\ C^\circ$ and at four different concentrations, namely 7.2 mg/ml, 12.2 mg/ml, 16.3 mg/ml and 20.0 mg/ml, which correspond to box sizes of, respectively, $L = 114$, 95.5, 86.4 and 81.1 nm. Fig. 5.2 shows simulation snapshots at different state points.

In the following analysis, we consider two tetramers as bonded if they share at least three bonded nucleotides. We note that changing this threshold does not significantly affect the results since nearly all of the bonded sticky ends are fully hybridised, i.e. all six available bonds are formed. This is to be expected, since free-energy profiles of duplex formation below the melting temperature show that full-hybridised structures are by far the most stable configurations if, as in the present case, the fraying effect is not relevant [110, 227].

One way of producing very low-density open equilibrium networks, the so-called empty liquids [39, 199, 205], is to employ particles that fulfil the single-bond-per-patch condition and do not form multiple bonds [38]. For the system studied in this Chapter the very nature of DNA hybridisation makes the former condition satisfied, while the large flexibility of the arms may favour multiple bonding between neighbours. In order to check if this is the case we computed the fraction of multiple bonds. The results show that, at the lowest studied temperature, no more than 5% of the bonds are multiple. Such a value should not have any qualitative effect on the network topology of the resulting gel.

In the following, if not otherwise stated, we use the centres of mass of

tetramers to carry out analyses of the configurations.

5.3 Results

5.3.1 Cluster size distribution

To investigate the static structure of the systems, we start by quantifying the extent of the self-assembling process. In order to do so we compute the cluster size distribution $n(s)$, i.e. the number of clusters of size s averaged over all the equilibrium configurations.

Fig. 5.3 shows the cluster size distributions for all the investigated systems. Despite the noisiness of the curves, due to the low number of simulated tetramers, there are some trends that are worth discussing. The number of monomers in the system is monotonically decreasing as the temperature is lowered, since more and more tetramers bind together. Even at the highest T, clusters made up of at least 20 tetramers are always present and when the temperature is decreased the average cluster size, shown in Fig. 5.4(a), begins to grow and, eventually, at the lowest T a percolating network of tetramers, which comprises most of the particles, is formed (disconnected points in the figure).

Fig. 5.4(b) shows the bond probability p_b, i.e. the probability that an arm of a tetramer is engaged in a bond. p_b, as expected [37], increases monotonically as the concentration is increased or the temperature is decreased.

We know from theoretical and numerical results that tetravalent patchy particle systems at the critical point have $p_b \approx 0.65$ [18]. On the other hand, since percolation is a prerequisite of second-order phase transitions [187–189], the percolation *locus* is encountered for lower values of the bond probability. For colloids coated with four short DNA strands, for example, p_b at percolation ranges from $p_b \sim 0.32$ to $p_b \sim 0.42$, depending on concentration [218].

5.3.2 Structure factor

In order to give an estimate of the location of the phase-separation region we study how the static structure factor $S(q)$ evolves with concentration and temperature. The results are shown in Fig. 5.5.

Upon lowering T, the $S(q)$ becomes more structured, with the height of the peaks and the depth of the minima increasing. The nearest-neighbour peak of the $S(q)$, related to the inter-tetramer bonding, is always located around $q \approx 0.5$, but its position moves slightly towards larger q as the concentration is

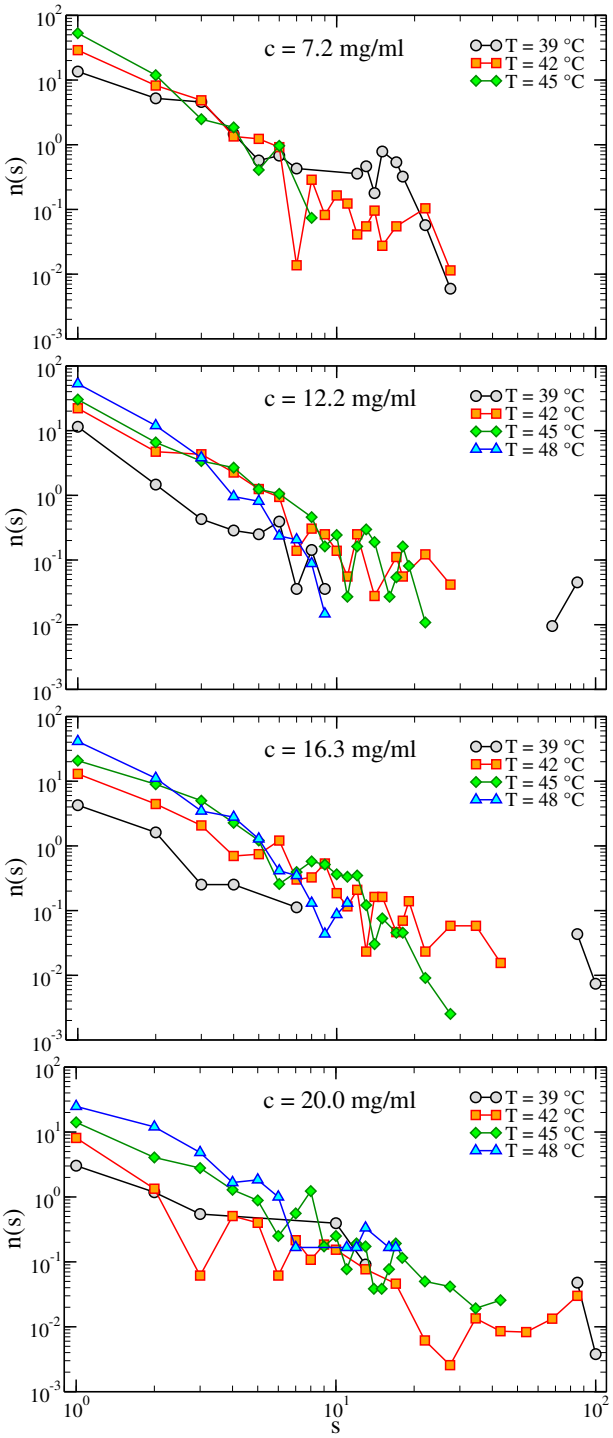

Figure 5.3: Cluster size distribution for all the investigated systems at fixed concentration and different temperatures. From top to bottom: $c = 7.2$ mg/ml, 12.2 mg/ml, 16.3 mg/ml and 20.0 mg/ml. Lines are guides for the eye.

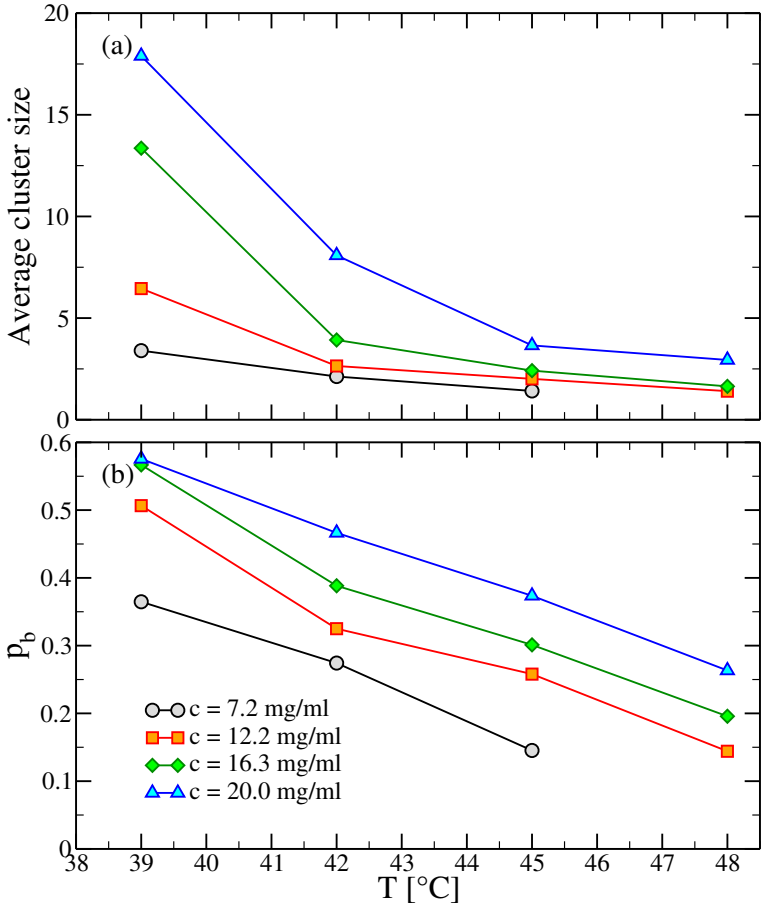

Figure 5.4: (a) Average cluster size and (b) bond probability p_b for all the investigated state points.

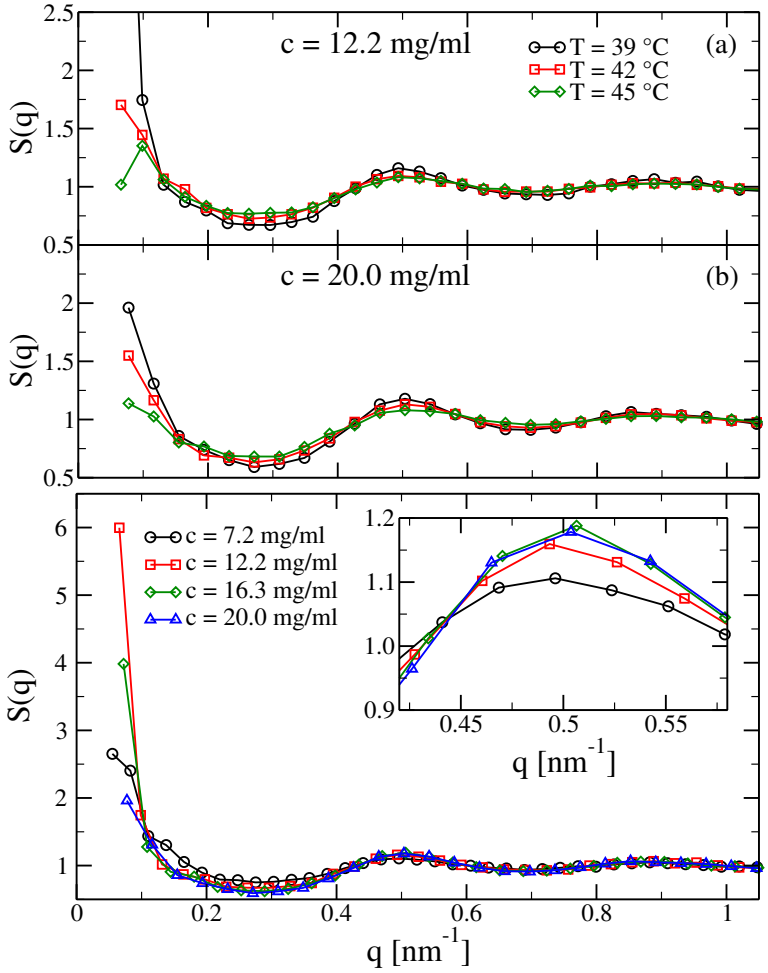

Figure 5.5: Top: static structure factor $S(q)$ for different temperatures for (a) $c =$ 12.2 mg/ml and (b) 20.0 mg/ml. The lowest-q value for the $T = 39$ °C, $c =$ 12.2 mg/ml $S(q)$ is 6 (not shown). Bottom: $S(q)$ at $T = 39$ °C for all the studied concentrations. The inset shows the first peak in greater detail.

increased. This value of q corresponds to a real-space distance of $r \approx 12.7$ nm. For comparison, the value obtained by considering two bonded tetramers in the minimum of the energy is $r \approx 14.8$ nm. This difference is probably due to the presence of the spacer separating the tetramer double-stranded arm and the sticky ends and to the non-rigid nature of dsDNA itself. Indeed, the persistence length of dsDNA is $\approx 150\ bp$, a value comparable with the number of base pairs forming two bonded tetramer arms, which is 48. The bending of the arms is evident in Fig. 5.1

On passing we note that, apart from the low-q limit, which will be discussed in the next paragraph, the structure of the systems at $c = 16.3$ mg/ml and $c = 20.0$ mg/ml at $T = 39\ ^\circ C$ is identical within numerical noise.

The growth of the low-q limit of the $S(q)$, related to the isothermal compressibility of the system via the relation $S(q \to 0) \propto \chi^{-1}$, signals an heterogeneity in the structure which could be due to the approaching of the gas-liquid phase coexistence region as well as to to the open nature of the network of tetramers. Indeed, the structure factors of equilibrium gels have non-negligible values of $S(q \to 0)$ which tend to saturate as, since at low T all the bonds are formed, the topology of the network does not evolve any more [3,65]. While the increasing of $S(q)$ for $c < 20.0$ mg/ml seems to indicate a phase separation, the lower value, and the lower rate of growth, of the low-q limit of the $S(q)$ of the highest concentration system may be a signal of the approaching to an equilibrium network structure [3,205]. It is also worth noting that the $c = 12$ mg/ml system has the highest $S(q)$, which may be a signal of the proximity of a critical point. Visual inspection of the configurations shows the inhomogeneous nature of the system (see Fig. 5.2).

The low-q limit of the structure factor can be used to give an estimation of the spinodal temperature T_s. Indeed, near a gas-liquid phase boundary the compressibility factor χ diverges as a power law with exponent $\gamma = 1.25$, and therefore

$$S(q \to 0)^{-1} \propto (T - T_s)^{-\gamma}. \tag{5.1}$$

We first extrapolate the values $S(q \to 0)$ using the low-q part of the structure factor by fitting it with a Lorentzian. Then we use the lowest-T values to extract the temperature at which the compressibility diverges via Eq. (5.1). Unfortunately, we can only use the $2 - 3$ lowest wave-vectors for extrapolation, as only these values correspond to the long length scales required. Fig. 5.6 shows

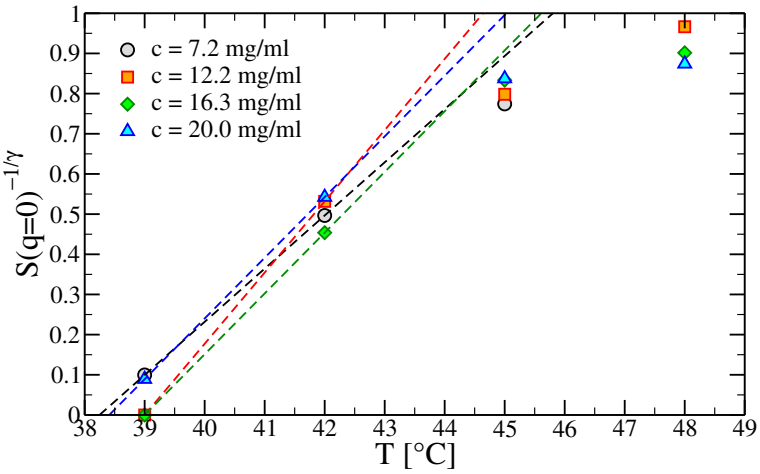

Figure 5.6: Extrapolated scattered intensity to the power of γ as a function of T for all the studied concentrations (points). Linearly extrapolating to 0 (dashed lines) gives the temperatures at which the spinodal is encountered for each value of c.

the numerical results and the obtained fitting curve for all the investigated concentrations. The complete phase diagram is presented in Fig. 5.10. All the extrapolations yield very similar results and hence the phase-coexistence region extends beyond $c = 20$ mg/ml. On the other hand, since concentrations of the liquid phase of limited-valence colloids are nearly independent of the temperatures [39], the spinodal curve always bends down really fast. Since the structure factor of the highest-concentration system does not reach very high values, we speculate that the phase-coexisting limiting concentration is located at concentrations slight larger than, but comparable with, $c = 20$ mg/ml. Simulations of higher concentration systems are under way.

5.3.3 Percolation

Now we compute the percolation probability p_p, i.e. the probability that the system contains a spanning cluster. Fig. 5.7 shows p_p against T for different values of the concentration.

At high temperatures there are no percolating clusters and, as shown in Fig. 5.3, the system is made up of small clusters which, upon lowering T, start to grow and eventually form a percolating network which spans the entire simulation box. This process happens at lower and lower temperatures as concentration is decreased. At the lowest studied concentration no percolation occurs. The percolation probability in finite systems has a sigmoidal shape that gets steeper as the system size is increased, eventually becoming a Heaviside

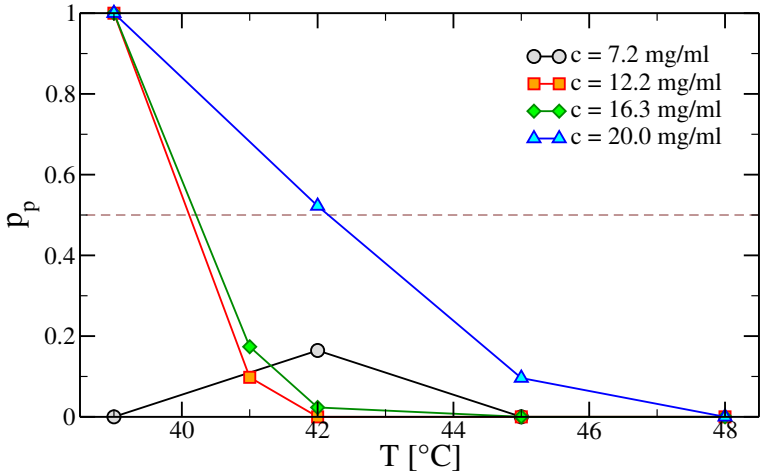

Figure 5.7: Probability percolation for all the investigated concentrations as a function of temperature. The $c = 7.2$ mg/ml, $T = 42°$C value is due to a percolating chain and therefore it is a finite-size effect.

step function in the thermodynamic limit. Since the computational cost of a finite-size study is prohibitive, we rely on the results obtained at a single box size. We estimate the percolation temperature T_{perc} by making cuts at $p_p = 0.5$ in Fig. 5.7. Fitting p_p with a sigmoidal and then considering again the temperature at which $p_p = 0.5$ results in very similar values of T_{perc}. The results are shown in Fig. 5.10.

Since the percolation line always intersects the spinodal *locus* at a concentration smaller than or equal to the critical concentration c_c, we can estimate the location of the latter in the range 10 mg/ml $\lesssim c_c \lesssim$ 15 mg/ml.

5.3.4 Mean-square displacement

Next we investigate the dynamics of the system by evaluating the mobility of the tetramers. Fig. 5.8 shows the mean-square displacement (MSD) $\langle r^2(t)\rangle$ for all the concentrations and $T \leq 45$ °C. At very short times the system follows an inertial dynamics and hence the MSD is proportional to t^2. After this ballistic regime, the system enters the diffusive regime, where $\langle r^2(t)\rangle \propto t$. At low temperatures an intermediate sub-diffusive regime, which is more and more pronounced as concentration is increased, appears. In the diffusive regime, the dynamics is controlled by the diffusion coefficient D, given by

$$D = \lim_{t\to\infty} \frac{\langle r^2(t)\rangle}{6t}. \tag{5.2}$$

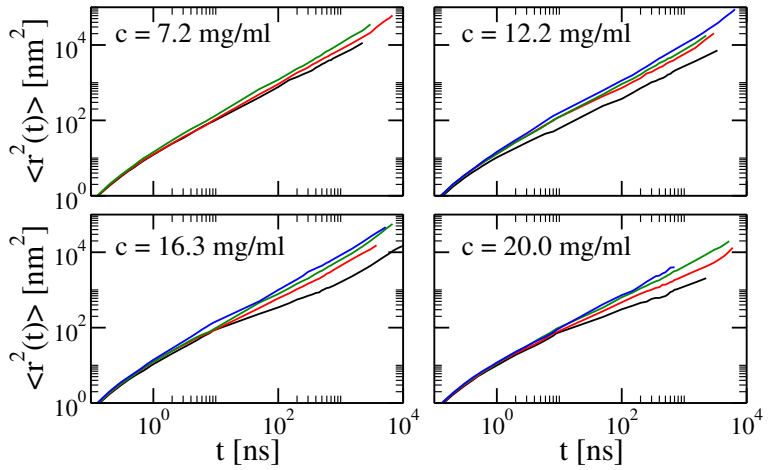

Figure 5.8: Mean-square displacement at $T = 39\ ^{\circ}C$ (black lines), $42\ ^{\circ}C$ (red lines), $45\ ^{\circ}C$ (green lines) and $48\ ^{\circ}C$ (blue lines) for all the studied concentrations.

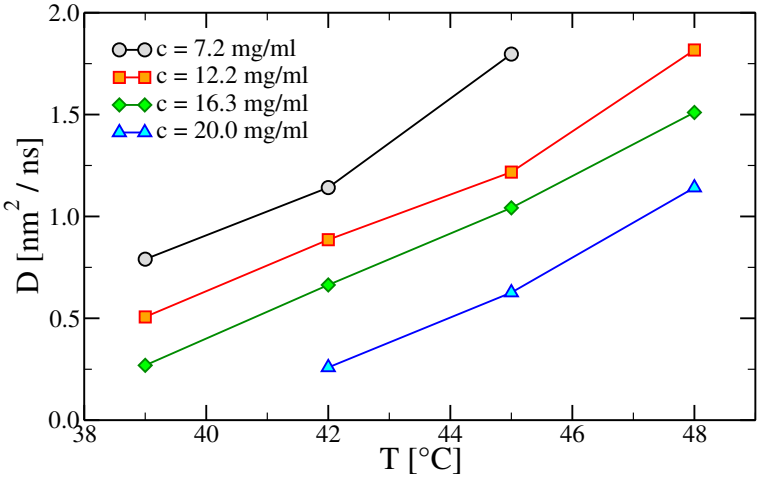

Figure 5.9: Diffusion coefficients for all the investigated state points. Lines are guides for the eye.

Fig. 5.9 shows D for all the investigated state points but the $c = 20$ mg/ml, $T = 39\ ^{\circ}C$ system whose MSD, as shown in Fig. 5.8, is still sub-diffusive. By making *cuts* at fixed values of D we obtain two isodiffusivity *locii* on the phase diagram, shown in Fig. 5.10.

5.3.5 Phase diagram

Fig. 5.10 shows all the investigated state points and the obtained results, as well as the experimental phase diagram recently measured by the group of Prof. Bellini in Milan [225]. The experiments have been carried out at the very low salt concentration 0.05 M, which explains the difference in temperature between the two phase diagrams [228, 229]. The difference in the concentration range, on the other hand, can not be explained by the ionic strength alone. Since the coarse-grained model we use has been parametrised to reproduce the structural properties of DNA, such a large error in the estimation of the coexisting concentration is not to be expected. A lower salt concentration may, in principle, lead to a lower percentage of double bonds since the repulsive interactions between the arms, negatively charged, would be higher and hence the probability of forming a double bond between tetramers having far apart arms would decrease. Since only a tiny fraction of the bonds are double, the net effect of this increased repulsion should not play a significant role in the concentration of the liquid branch.

The observed concentration difference could be tentatively ascribed to the incomplete formation of the tetramers in the experimental systems for two different reasons. First, the number of constituent strands is not exactly the same, and therefore not every strand can be part of a tetramer. Secondly, it has been shown that the yield of complex clusters in bulk systems has a very broad transition as a function of temperature, and therefore the annealing procedure will always leave some structures unformed [230]. Thus the effective valence of the system, which depends on the number of single strands, dimers and trimers, will be always lower than 4, and hence the measured spinodal curve will be moved at lower temperatures and concentrations with respect to the $M = 4$ system.

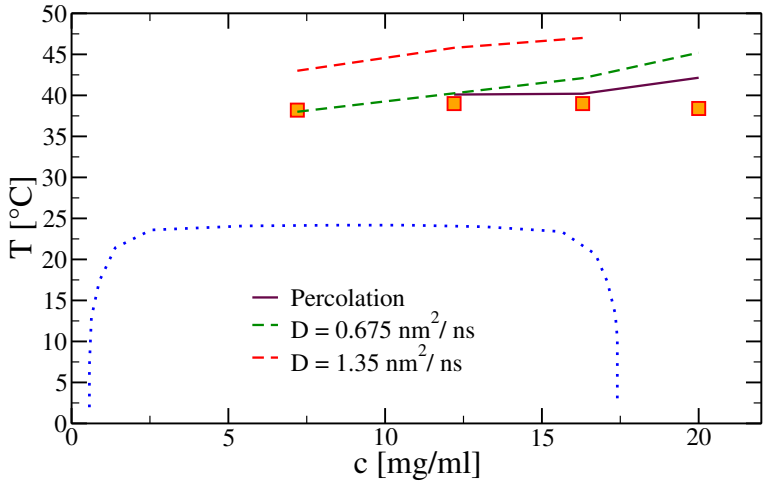

Figure 5.10: Phase diagram of the system. By extrapolating the scattered intensity of all the investigated state points we are able to estimate the location of the spinodal curve (red squares). The blue dotted line is the experimental phase diagram measured by the group of Prof. Bellini in Milan [225]. Dashed lines are isodiffusivity curves, while the solid line is the percolation *locus*.

5.4 Conclusions

In this Chapter we have investigated the phase behaviour and the dynamics of DNA tetramers, i.e. DNA constructs of valence $M = 4$. The primary constituents are ssDNA molecules which have been designed to first self-assemble into tetramers and then, upon lowering the temperature, to reversibly form networks of inter-tetramer bonds. This experimental realisation of limited-valence particles is expected to undergo at low temperatures a phase separation between a gas-like diluted phase and a low-concentration, percolating, liquid-like phase. This opens up a region of intermediate concentrations in which the system can be cooled down without encountering any thermodynamic instability. By changing the valence, i.e. by employing constructs with a smaller number of arms, the properties of the resulting thermoreversible gel, an empty liquid [39, 205], can be opportunely tuned.

The present study, featuring very lengthy, large-scale numerical simulations on GPUs, is a step forward in the direction of quantitatively predicting the whole phase diagram of DNA constructs. Indeed, being able to calculate the thermodynamic and dynamic behaviour of complex systems is of paramount importance for designing the materials of tomorrow. By using a sophisticated, realistic model of DNA, simulated on GPUs in order to obtain the required performance speed-up, we have managed, for the first time, to study a bulk

system composed of 19600 nucleotides. We have observed the formation of clusters and, eventually, the appearance of a inhomogeneous percolating network signalling the approaching of the phase separation.

The calculated phase diagram bears some similarities with very recent experimental results by the group of Prof. Bellini [225]. Unfortunately, the experiments were performed at a very different salt concentration, making it impossible to quantitatively compare the two phase diagrams. Nevertheless, some preliminary results shows that carrying out experiments at a salt concentration of 0.5 M increases the spinodal temperature by $10 - 15\ ^{\circ}C$, values which are in line with numerical results. We have also shown that it is possible to compute also some characteristic *locii* like percolation lines and isodiffusivity curves.

In this Chapter we have shown that it is indeed possible to compute the thermodynamics and dynamic properties of DNA constructs, providing a way of helping in designing novel materials with tunable properties. The reported simulations partly suffer from size limits but show a lot of potential for future applications. Indeed, with the computational power ever increasing, it will be soon possible to perform studies like the present one in a very automated and fast way, so that it will be possible to rely on these results to finely adjust the material properties.

Chapter 6

GPU programming: molecular dynamics on steroids

6.1 Introduction

For fifty years the performances of computing machines have been increasing steadily at a very high rate, with the number of transistors per CPU doubling every 18–24 months (the so-called *Moore's law*). As a result, over the last decades computer simulations have flourished as valuable tools to close the gap between theory and experiment, as they can be used to both test theoretical results as well as to interpret and predict experimental outcomes. With the increase of available computational power, more complex (i.e. larger or more realistic) systems can be simulated. As a step forward in this direction, a lot of effort has been recently put into the design and production of programmable Graphics Processing Units (GPUs) which were originally developed as specialised hardware for 3D rendering in computer games and industrial applications. As a consequence of this specialisation, the architecture of modern GPUs is intrinsically parallel, with hundreds of streaming processors implementing the single-instruction-multiple-threads (SIMT) paradigm and peak performances which are one or even two orders of magnitudes higher than those of CPUs. Thanks to new developed technologies, such as NVIDIA's CUDA [204], it is now possible to exploit this huge computational horsepower for general purpose tasks.

In Section 6.2 we give a brief overview of the architecture of a modern NVIDIA GPU and then, in Section 6.3, we introduce the algorithms and tech-

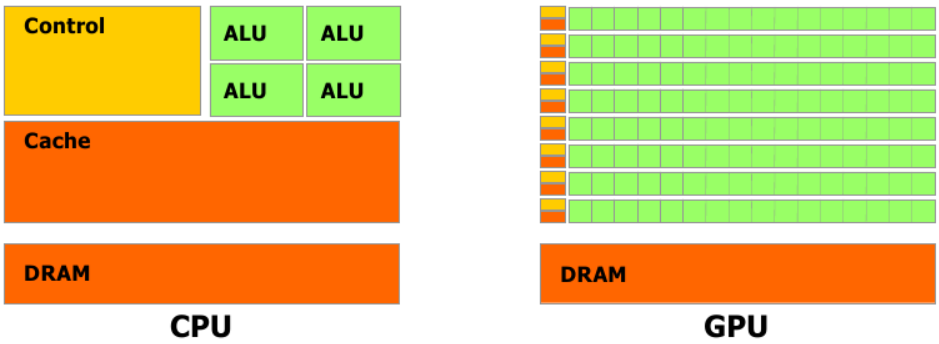

Figure 6.1: (Reproduced from Ref. [231]). Schematic comparison between CPU (left) and GPU (right) architectures. In GPUs, more transistors are devoted to number-crunching rather than flow control and data caching. ALUs are arithmetic logic units, i.e. digital circuits that performs arithmetic and logical operations. DRAM is the memory of the system.

niques employed to implement a general molecular dynamics code on GPU. Section 6.4 presents results on the performances of the code under different conditions (number of particles, density, temperature, simulated model).

6.2 GPU architecture

In this Section we will give a concise overview on the architectural features of NVIDIA GPUs. For a more complete description we refer the reader to the official CUDA programming guide [231].

Modern GPUs are specialised for compute-intensive, highly parallel tasks, such as graphics rendering. Compared to CPUs, whose more general role implies that a significant part of their transistors is devoted to flow control, branch prediction and data caching, GPUs are much more data-processing oriented, as shown in Fig. 6.1.

The most important feature of NVIDIA GPUs is the SIMT model, which makes these devices particularly well-suited to perform the same operations on different data. This implies that, in order to obtain the best results (e.g., high memory throughput or near-to-close GFLOPs peak performances) one has to fully understand both the hardware architecture and the software abstraction layer built on top of it.

Designed to provide a computing architecture that makes use of the parallel compute engine in NVIDIA GPUs to solve parallel, complex computational problems in a very efficient way, NVIDIA's CUDA includes a programming

environment that allows the developing of applications which utilise GPUs by extending the standard C/C++ syntax.

The software layer used by CUDA to hide the hardware implementation, so that GPU programs can be run on any suitable device, is based on three fundamental concepts:

- kernels
- thread hierarchy
- memory hierarchy

In the next paragraphs we will briefly analyse each point.

6.2.1 Kernels

Functions executed on the GPU are called *kernels*. When called, a kernel is executed N times in parallel by N different threads, with each thread having a unique *thread ID* which unambiguously identifies it within a block of threads. Kernels are identified by the keyword `__global__`. The number of threads, their hierarchy and the amount of required shared memory are specified at runtime through a special *execution configuration* syntax.

6.2.2 Thread hierarchy

Threads executing a kernel are divided into *blocks* of threads whose dimension is chosen at runtime. Blocks can be one-dimensional, two-dimensional or three-dimensional, providing a natural way of mapping elements of, respectively, vectors, matrices or volumes. All threads within a block reside on the same multi-processor (MP) and thus have shared access to a part of the memory resources of that MP, and therefore can communicate with each other using primitive synchronisation barriers. Blocks are then organized into a one-, two- or three-dimensional *grid* of blocks. This further division is practical rather than logical because, since threads of different blocks have no primitive way of communicating with each other, using a multidimensional rather than a unidimensional grid has the only advantage of providing a simpler way to access multidimensional data structures.

6.2.3 Memory hierarchy

For both architectural and practical reasons, CUDA memory is organised in different levels, each with its own advantages and drawbacks. As a general rule, faster memories have a smaller size. The most important types of memory are:

1. Global memory is the slowest but also the most abundant memory (few GBs on most modern devices). It has the lifetime of the application and it can be accessed by all threads. Performances can be greatly improved by coalescing memory accesses, i.e. by making threads which have consecutive IDs access contiguous memory regions. Concurrent accesses can be regulated using atomic functions [231].

2. Constant memory is a read-only memory which resides in global memory. It is very limited in size (tens of KB) and it has the lifetime of the application. It is cached, and therefore is used to store constant parameters used throughout the simulation like temperature, box size, number of cells, etc.

3. Shared memory is a very fast (two order of magnitudes faster than global memory) but low-capacity memory visible to all the threads of the block, and hence has the lifetime of the block. If used correctly it can greatly enhance performances. It is commonly used to avoid unnecessary memory reads by storing values which reside in global memory and are required by all threads in a block, so that the values are fetched from global memory only once per block.

4. Local memory is a per-thread memory used to store all the local variables. If the required memory size is not too high, the compiler places all the variables in the registers, but if this is not the case then global memory is used (the so-called *register spilling*), thus vastly reducing performances. It is very important to makes kernels use as small amount of memory as it is possible in order to optimize memory throughput.

6.2.4 Conclusions

All these features, which are simply exposed to the developer as a minimal set of language extensions, provide fine-grained data parallelism and thread parallelism, nested within coarse-grained data parallelism and task parallelism. They guide the programmer to partition the problem into coarse sub-problems

that can be solved independently in parallel by blocks of threads, and each sub-problem into finer pieces that can be solved cooperatively in parallel by all threads within the block. This decomposition allows for both compatibility and scalability: since only threads in the same block can cooperate when solving each sub-problem, each block of threads, which hence does not depend on any other block, can be scheduled on any of the available multiprocessors within a GPU, in any order, concurrently or sequentially, so that a compiled CUDA program can execute on any number of multi-processors. Since this process is completely transparent to the user, only the runtime system, which handles the scheduling, needs to know the physical multiprocessor count [231].

6.3 Molecular dynamics on GPU: an implementation

In this Section we describe the implementation of the MD code used in Chapters 1, 4 and 5 to study the thermodynamics and dynamics of patchy particles and DNA constructs.

Unlike most types of Monte Carlo simulations, MD simulations are naturally suitable for parallel architectures since they perform the same operations on each particle. Even though there have been some successful attempts at parallelising lattice Monte Carlo simulations [232, 233], the intrinsic serial nature of off-lattice MC simulations hinders the possibility of high performance gains. Indeed, in MD simulations the integration steps are straightforwardly parallel because each thread, which performs the integration in time of a single particle's degrees of freedom and momenta, is independent and executes without the need of communicating with any other thread. The calculation of forces and torques, on the other hand, requires a careful design in order to minimise fetches from global memory. In the next paragraphs we will discuss

1. how to implement cells and Verlet lists, so that the computational complexity of an N-body simulation is $\mathcal{O}(N)$.

2. How to speed-up memory accesses by sorting particles so that positions and momenta of neighbouring particles are stored in adjacent memory regions.

3. What is, performance-wise, the most convenient way of retaining a sufficiently high precision without having to resort to a full double-precision simulation, since double-precision floating-point operations halve the memory bandwidth and are a lot less efficient than single-precision ones.

6.3.1 Cells and Verlet lists

Since, in principle, the force acting on particle i depends on all the other N particles, the computational complexity of a MD code is $\mathcal{O}(N^2)$. Nevertheless, if the interaction potential acting between particles has a cut-off radius r_c and the simulation box is divided into cubic cells of size $r_{\text{cell}} \geq r_c$, only particles in neighbouring cells will feel the mutual interaction. Since the number of neighbouring cells does not depend on the total volume, computing the forces acting on each particle is a $\mathcal{O}(1)$ operation, which means that the computational complexity of computing the interactions between all the particles, and hence of the whole simulation program, is $\mathcal{O}(N)$. Even if the most common way of dealing with cells is to use the linked-list method, which allows to both save memory and keep the list of particles in each cell continuously updated [172], implementing this method on GPUs is not feasible since keeping track of particle displacements between cells is an intrinsic serial operation. To overcome this limitation we build the cell lists only when Verlet lists have to be updated. Cell lists are stored in a $N_c \times n_{\text{max}}$ matrix, with N_c being the number of cells and n_{max} the maximum number of particles in a cell, to be estimated at the beginning of the simulation. The number of particles in each cell is stored in an array of size N_c. The implementation is as follows: each thread takes care of a single particle and updates both the cell matrix and array by computing the index of the particle's cell and updating the data structures accordingly. This is done using atomic functions to avoid race conditions. It is also possible to come out with a more sophisticated algorithm that makes use of shared memory, but this is convenient only for intermediate–high densities ($\rho \geq 0.3$) [234].

Verlet lists, which are built using cell lists as previously stated, are used to further lower the number of neighbours per particle. The implementation is very similar to what is done on CPUs [172], with each thread i building the neighbouring list of the i-th particle by looping over all the particles that are in one of the 27 neighbour cells of that particle.

6.3.2 Particle sorting

As stated in Section 6.2.3, memory performances are maximised when reads and writes are performed in a coalesced way. In a MD code, since the data structures of spatially neighbouring particles are, in general, stored in non-contiguous memory regions, i.e. randomly, threads executing the interaction-computing kernel access memory in a non-coalesced way when fetching the

data structures relative to a particle's neighbours. The resulting heavy memory overhead is reduced if the data is somehow cached, since fetching a value from the cache is order of magnitudes faster than re-reading it from global memory. This is straightforward on newer architectures since, starting with devices of compute capability 2.x, NVIDIA GPUs feature a unified L2 cache (i.e. a cache shared by all multiprocessors), while older devices require the usage of textures in order to have a reasonably sized cached memory. To increase the number of *cache hits* it is convenient to periodically sort all the particles, and their data structures, according to a three-dimensional space-filling curve. With this method, introduced by Anderson *et al.* [234], each particle is mapped onto a Hilbert curve and it is given an index relative to its one-dimensional position on the curve. This index is then used to permute all positions, orientations and momenta so that neighbouring particles in space have their data structures stored in nearby memory regions. This 3D to 1D mapping can vastly increase performances for large systems, as shown in Section 6.4.

6.3.3 Single vs double precision

For the integration of the equations of motions we adopt the velocity-Verlet algorithm which, being symplectic, conserves the average internal energy $\bar{U}$ if no thermostat is employed, that is, if the simulation is performed in the NVU ensemble [235]. The fluctuations of $U(t)$ around $\bar{U}$ are related to the integration step δt via

$$\sigma_U = \left\langle (U(t) - \bar{U})^2 \right\rangle \propto \delta t^2. \tag{6.1}$$

If δt is too large or if it is of the same order of magnitude of the precision of the simulation, Eq. (6.1) is not valid any more and the energy starts to drift away from its average value. If this is the case then the simulation is *numerically unstable* and its results are unreliable. As a general rule, the steeper is the potential, the smaller δt has to be in order to maintain numerical stability. Of course, a smaller δt leads to longer simulations since more MD steps are needed to simulate a given time interval. Care must be taken in order to choose a value of δt which is as large as possible but does not result in any energy drift on the timescale of the simulation. For regular, double-precision (DP) simulations on CPUs involving common potentials (e.g. Lennard-Jones, generalised Lennard-Jones, FENE, harmonic, gaussian potentials), $\delta t = \mathcal{O}(10^{-2}) \div \mathcal{O}(10^{-4})$ in simulation units.

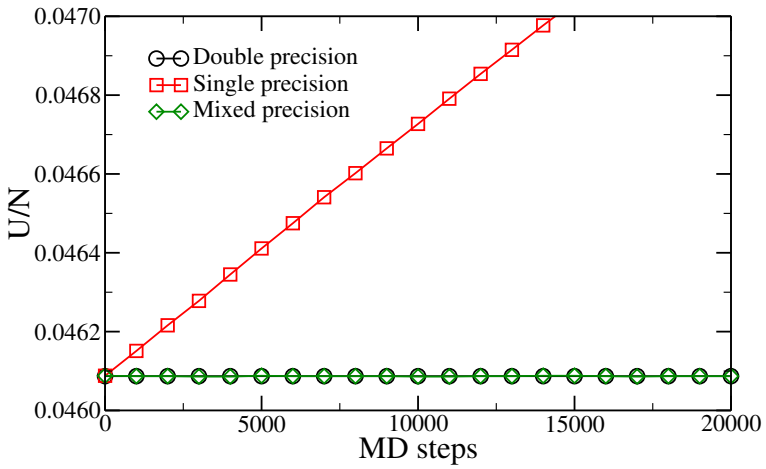

Figure 6.2: Energy per particle as a function of simulation time (expressed in MD steps) for a patchy-particle system. Data for double-, single- and mixed-precision simulations with $\delta t = 10^{-4}$ in the NVU ensemble are reported. On the scale of the plot, the mixed- and double-precision curves fall on top of each other.

Early NVIDIA GPUs did not support DP floating-point operations, since rendering applications rarely need such a high precision. Starting with devices of compute capability $1.x$, hardware support for DP floating-point operations was added. Even if DP efficiency is being increased with every new GPU generation, employing DP calculations on GPUs still has two main drawbacks:

- The number of DP floating-point computational units is less than its single-precision (SP) counterpart. This implies that performing DP calculations is not as efficient as performing SP calculations.
- On most architectures, CUDA being one of them, a DP floating-point number takes up 64 bit, doubling the amount of memory required by SP floating-point numbers. Therefore, using DP calculations results in an increased register usage and in worse overall memory performances (see Section 6.2.3). Memory-bound applications like MD simulations are deeply affected by this factor.

Unfortunately, it has been shown that lengthy SP simulations lack reliability even if simple potentials are employed [236]. In the case of very steep potentials, such as the one we use to perform simulations of patchy particles [40, 65], even short simulations are hindered by the single precision. As an example, the difference in the conservation of the total energy between SP and DP simulations of a patchy-particle system is shown in Fig. 6.2. Even with a small integration

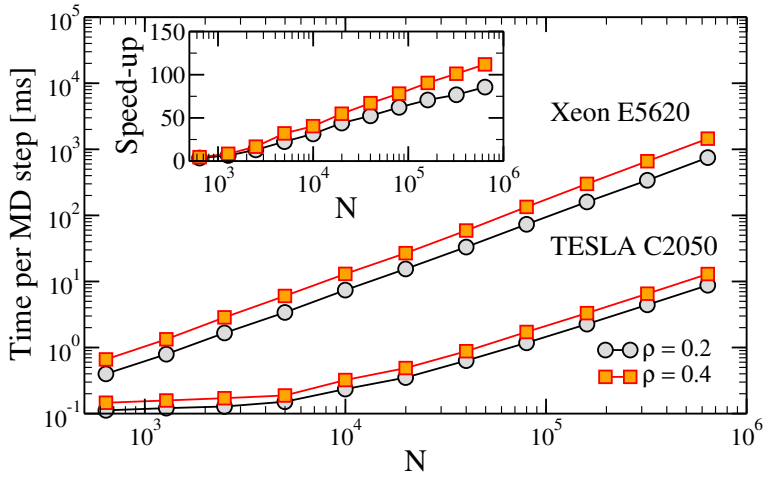

Figure 6.3: Total time per MD step against number of particles N for two densities, $\rho = 0.2$ and $\rho = 0.4$ and for CPU (top curves) and GPU (bottom curves) in the case of a Lennard-Jones potential. Inset: achieved speed-ups on GPUs.

time step the SP simulation is clearly numerically unstable: the energy starts drifting at the very beginning of the simulation. In the DP case, on the other hand, the energy fluctuates around its average value with $\sigma_U \sim 10^{-9}$, but the simulation, tested on a NVIDIA C2050 device, is 8 times slower than the SP one.

Since full-DP simulations are still up to an order of magnitude slower than SP simulations, one has to limit the usage of DP calculations to the portion of the code where a higher precision is critical. Careful analyses show that the most critical part, with respect to numerical stability, is the integration of coordinates and momenta, which we implement in double precision. The most time-consuming part of the simulation code, the evaluation of the forces, is carried out in single precision. This mixed-precision (MP) method results in a decrease of performances comprised between 10% and 40%, depending on model, state point and number of particles. Result-wise, the green curve in Fig. 6.2 shows that MP simulations are numerically stable when SP simulations are not. For a thorough discussion on this topic we refer the reader to Ref. [236].

6.4 Performances

In this Section we analyse the performances of the developed MD code for systems interacting through a simple Lennard-Jones potential and through a coarse-grained, realistic DNA potential [110].

Fig. 6.3 shows the achieved performances and speed-ups for different num-

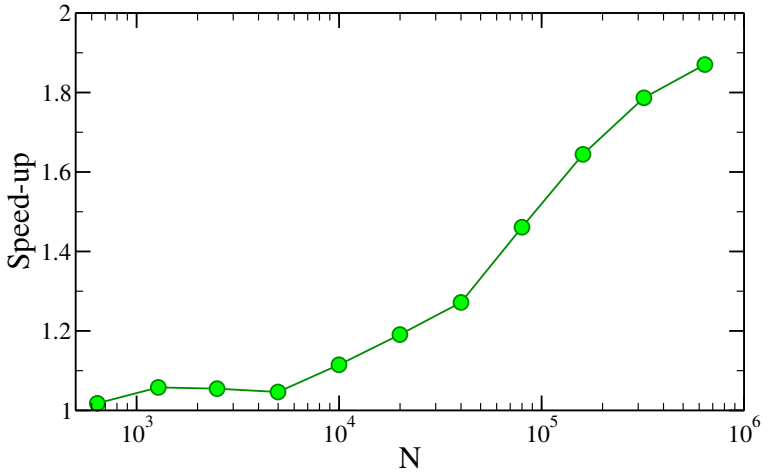

Figure 6.4: Speed-ups achieved employing particle sorting (see Section 6.3.2) for different numbers of particles at fixed $\rho = 0.2$.

bers of particles at two fixed densities for Lennard-Jones systems. The time step used was $\delta t = 10^{-4}$ and the Verlet skin for both CPU and GPU was set to 0.3. In both cases the linear regime for GPU simulations kicks in at approximately 10^4 particles: for smaller values of N, simulations are memory bound and therefore simulating systems with less particles is not convenient. As shown in the inset, speed-ups of two orders of magnitude can be achieved for very large systems, although precise performances depend on simulation details, in particular on density, time step and Verlet skin values.

For very large numbers of particles the negative effects of randomly accessing memory becomes prominent. Fig. 6.4 shows that sorting particles data, as described in Section 6.3.2, is very effective and provides a significant boost of performances for $N > 10^5$.

Because of the GPU SIMT architecture, the performances achieved on graphical devices can be hindered by complicated potentials involving many conditional statements, e.g. piecewise functions. Nonetheless, GPUs can still be successfully employed to simulate these systems. In particular, we implement the coarse-grained DNA model developed by Ouldridge *et al.* [110, 112], which we then employ to study bulk systems in Chapters 1 and 5. In this model, nucleotides interact through seven non-trivial, short-ranged potentials. In turn, each of these potentials is made up by piecewise functions, made continuous by a gaussian smoothing.

Fig. 6.5 shows the performances obtained for a system containing 20-base-pairs duplexes. In this case, as in Lennard-Jones simulations, the time per MD

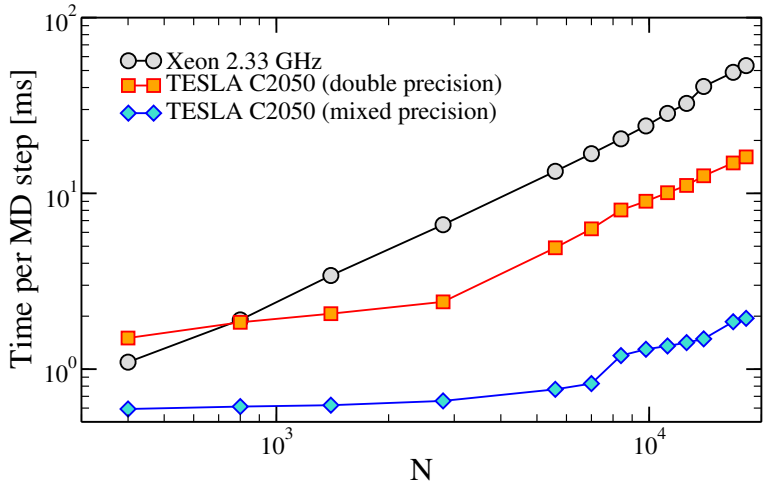

Figure 6.5: Total time per MD step against number of nucleotides N for systems formed by 20-base-pairs long duplexes at very low concentrations. Result obtained with CPU (grey circles), DP GPU (orange squares) and MP GPU (cyan diamonds) simulations.

step starts to behave linearly only for $N \gtrsim 10^4$. In this plot we also compare the performances of DP versus MP simulations on GPUs: as stated in Section 6.3.3, DP simulations are up to an order of magnitude slower than MP simulations. It is worth noting that the Ouldridge *et al.* model, being highly branched, is approximately two times more resource-demanding than a simple Lennard-Jones system on CPU, whereas on GPUs the difference is almost an order of magnitude.

The observed speed-ups for DNA systems with $\sim 10^4$ range from 20 to 40, depending on the simulated system. Given the high degree of branching of the potential, i.e. the number of possible branches in the code due to conditional statements, GPUs are more affected by the number of neighbours than CPUs and hence speed-ups and performances heavily depend on the investigated system.

Conclusions

One of the main goals of present day material science is to design materials with tunable properties by following a bottom-up approach, that is, by carefully adjusting the properties of the building blocks so that they self-assemble into the required target structures. However, in order to do so we have to be able to link the microscopic details of the interaction between the constituents to the macroscopic properties of the bulk system. Indeed, despite the spectacular advances achieved in this field in the last decade, the lack of a more general and unified point of view on the building blocks hinders further progresses [8]. As a step forward in this direction, in this Thesis we have investigated the bulk behaviour of anisotropically interacting colloidal particles by means of theoretical and numerical calculations, with a particular emphasis on the latter.

In Chapter 1 we study a dense solution of short DNA duplexes exhibiting an isotropic–nematic phase separation. These cylindrical particles, when the temperature is lowered or the concentration is increased, self-assemble into longer and longer chains by forming bonds due to hydrophobic forces. We combine a parameter-free theoretical approach with very large-scale numerical simulations of a coarse-grained, realistic DNA model [110, 112], to compute the location of the phase separation. We first test the theory in the isotropic phase by comparing its outcomes to numerical results, finding good agreement. Then we use it to predict the location of the IN phase boundaries for different duplex lengths. The comparison between theoretical and experimental results [49, 90, 91] is satisfactory. The joint approach we have developed to quantitatively predicting the IN phase boundary location of DNA duplexes can be simply extended to study the self-assembly of generic chain-forming anisotropic particles, and hence it is a valuable tool that can be used by theorists and experimentalists to both understand and predict the phase behaviour of nematic and cholesteric systems.

The dipolar hard-sphere (DHS) model is a paradigm for the self-assembly

of anisotropic particles and a challenge for present day theories of fluids. For these reasons, significant effort has been put into the direction of computing its phase diagram. In Chapter 3 we study the behaviour of the DHS model at low temperatures and low densities through extensive state-of-the-art numerical simulations, shedding light on a region of the phase diagram where a topological phase transition has long been thought to occur [55]. We show that the system exhibits remarkable and unusual behaviours, like a very low density percolation locus and a stabilization of rings over chain structures. This unexpected abundance of rings comes from a delicate balance between the lower ring energy and the end-to-end chain entropy, and hints at a possible mechanism for the suppression of the gas-liquid phase separation. Our results open the possibility for refined theoretical approaches which, in addition to the previously encompassed chain and branched geometries, must also include the significant contribution arising from ring formation. In Chapter 2, building on these results, we want to establish a clear link between patchy models and dipolar models, which are very demanding from both the numerical and the theoretical point of view. On the contrary, patchy models are much easier to be simulated and treated theoretically but, at the same time, can be designed to incorporate any kind of anisotropy by carefully choosing the number of patches, their shape and interaction strength.

By using an approach similar to what we have done in Chapter 1, we study a simple fluid composed of particles having a hard-core repulsion complemented by two patchy attractive sites on the particle poles. We choose the patch angular width to qualitatively reproduce the behaviour of very low-density DHS fluid, as reported in Chapter 3. Indeed, the chosen geometry favours the formation of ring structures which, at low temperatures and low densities, compete with the growth of linear aggregates. The simplicity of the model makes it possible to compare simulation results and theoretical predictions based on the Wertheim perturbation theory, specialized to the case in which ring formation is allowed. Such a comparison offers a unique framework for establishing the quality of the analytic predictions. We have found that the augmented Wertheim theory describes remarkably well the simulation results.

In Chapter 3 we also report accurate calculations of the DHS particle. Implementing efficient and tailored Monte Carlo algorithms, we are able to explore, in equilibrium, the low-temperature region where a phase separation between a dilute gas of chain ends and a high-density liquid of chain junctions has been predicted to occur. Our data clearly show that the density of states remains

always single peaked, definitively excluding the possibility of critical phenomena in the investigated temperature and density region. We speculate that the abundance of rings is responsible for the suppression of the phase separation. For this reason, we intend to extend the theory developed in Chapter 2 by adding support for branching in order to fully understand the phase behaviour of dipolar-like fluids. Such a work is under way.

The dynamic properties of a material are of fundamental importance for its technological applications. The properties of thermoreversible gels, being equilibrium systems, can be tuned by both designing *ad-hoc* microscopic constituents and by changing the external parameters, such as temperature or pressure. To gather further insights in this direction, in Chapter 4 we discuss the results of intensive Brownian dynamics GPU simulations of a simple model of tetrahedral patchy particles in the optimal network density region, i.e. where the system can be cooled down to form a fully bonded disordered network. This choice allows us to investigate the evolution of the structure and of the dynamics in a wide range of temperatures without encountering any phase separation or dynamic slow-downs due to excessive packing. We find that the observed slowing down of the dynamics in this model system is driven by the progressive bond formation and the increasing bond lifetime, which is controlled by temperature and interaction details. Although dynamical arrest is different from the glass case, where excluded volume interactions are dominant, the decay of the self- and collective correlation functions of the resulting fluid bears similarities with that observed in glassy systems.

In Chapter 5 we investigate a possible experimental realisation of a tetravalent system: tetramers made up of four single-stranded DNA molecules. This system is somewhat similar to the tetravalent patchy model studied in the previous Chapter and as such it is expected to exhibit a gas-liquid-like phase separation at low concentrations. At the optimal density region, this system should thus behave like an equilibrium gel, forming a stable, percolating network of bonds as the temperature is lowered. We use the coarse-grained, realistic DNA model presented in Chapter 1 to investigate the phase behaviour of this system with very large-scale numerical GPU simulations, equivalent to several decades of CPU time. As expected, the tetramers aggregate upon cooling and, if the concentration is sufficiently high, a percolating network appears. We calculate the location of the phase boundary, of the percolation line and of isodiffusivity curves.

Predicting in a quantitative way the equilibrium behaviour of realistic self-

assembling systems is a challenge to be taken on with sophisticated and up-to-date scientific tools. In Chapter 6 we introduce the techniques required to write an efficient Molecular Dynamics GPU code which makes it possible to simulate very large and complicated systems. Indeed, depending on the investigated model, performance boosts of 30x – 200x are achieved. The code presented in this Chapter has been extensively used to compute the results reported in Chapters 1, 4 and 5.

Our work shows that it is possible to predict the location of thermodynamic and dynamic *locii* of very complicated objects by means of numerical simulations. Since the available computational power keeps increasing at a steady pace, it will be soon possible to repeat the pioneering study presented in this Thesis on a more automated basis and for even more complicated system. For example, it will be possible to directly study the isotropic–nematic phase transition of short DNA duplexes investigated in Chapter 1 or design self-assembling DNA strands able to reproduce the behaviour of the patchy colloids or dipolar fluids studied throughout this Thesis.

Being able to carefully design the building blocks and then predict beforehand the properties of a compound will greatly simplify the process of synthesising tomorrow's materials.

List of publications

- L. Rovigatti, W. Kob, and F. Sciortino. "The vibrational density of states of a disordered gel model". *J. Chem. Phys.*, **135** (10), (2011) p. 104502.
- L. Rovigatti and F. Sciortino. "Self and collective correlation functions in a gel of tetrahedral patchy particles". *Mol. Phys.*, **109** (23-24), (2011) p. 2889.
- L. Rovigatti, J. Russo, and F. Sciortino. "No Evidence of Gas-Liquid Coexistence in Dipolar Hard Spheres". *Phys. Rev. Lett.*, **107**, (2011) p. 237801.
- L. Rovigatti, J. Russo, and F. Sciortino. "Structural properties of the dipolar hard-sphere fluid at low temperatures and densities". *Soft Matter*, **8**, (2012) p. 6310.
- C. De Michele, L. Rovigatti, T. Bellini, and F. Sciortino. "Self-assembly of short DNA duplexes: from a coarse-grained model to experiments through a theoretical link". *Soft Matter*, **8**, (2012) p. 8388.
- J. M. Tavares, L. Rovigatti, and F. Sciortino. "Quantitative description of the self-assembly of patchy particles into chains and rings". *J. Chem. Phys.*, **137** (4), (2012) p. 044901.
- P. Šulc, F. Romano, T. E. Ouldridge, L. Rovigatti, A. A. Louis, and J. P. K. Doye. "Introducing sequence-dependent interactions into a coarse-grained DNA model". *J. Chem. Phys.*, **137** (13), (2012) p. 135101.

Bibliography

[1] I. Hamley. *Introduction to Soft Matter*. Wiley & Sons, **2007**.

[2] S. C. Glotzer. "Some assembly required". *Science*, **306**, (2004) pp. 419–420.

[3] J. Russo, P. Tartaglia, and F. Sciortino. "Reversible gels of patchy particles: Role of the valence". *J. Chem. Phys.*, **131**, (2009) p. 014504.

[4] K. Neuman and A. Nagy. "Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy". *Nat. Methods*, **5** (6), (2008) pp. 491–505.

[5] M. Padgett and R. Di Leonardo. "Holographic optical tweezers and their relevance to lab on chip devices". *Lab Chip*, **11**, (2011) pp. 1196–1205.

[6] Y. Seol, G. M. Skinner, and K. Visscher. "Elastic Properties of a Single-Stranded Charged Homopolymeric Ribonucleotide". *Phys. Rev. Lett.*, **93**, (2004) p. 118102.

[7] Y. Seol, G. M. Skinner, K. Visscher, A. Buhot, and A. Halperin. "Stretching of Homopolymeric RNA Reveals Single-Stranded Helices and Base-Stacking". *Phys. Rev. Lett.*, **98**, (2007) p. 158103.

[8] S. C. Glotzer and M. J. Solomon. "Dimensions in anisotropy space: rationalizing building block complexity for assembly". *Nat. Mat.*, **6**, (2007) pp. 557–562.

[9] P. N. Pusey and W. van Megen. "Phase behaviour of concentrated suspensions of nearly hard colloidal spheres". *Nature*, **320**, (1986) p. 340.

[10] P. N. Pusey and W. van Megen. *Phys. Rev. Lett.*, **59**, (1987) p. 2083.

[11] S. Asakura and F. Oosawa. "Interaction between particles suspended in solutions of macromolecules". *Journal of Polymer Science*, **33** (126), (1958) pp. 183–192.

[12] C. N. Likos. "Effective interactions in soft condensed matter physics". *Physics Reports*, **348**, (2001) p. 267.

[13] J. P. Hansen and I. R. McDonald. *Theory of simple liquids*. 3rd edition. Academic Press, New York, **2006**.

[14] M. Muschol and F. Rosenberger. "Liquid–liquid phase separation in supersaturated lysozyme solutions and associated precipitate formation/crystallization". *J. Chem. Phys.*, **107** (6), (1997) p. 1953.

[15] P. J. Camp. "Phase diagrams of hard spheres with algebraic attractive interactions". *Phys. Rev. E*, **67**, (2003) p. 011503.

[16] M. H. J. Hagen and D. Frenkel. "Determination of phase diagrams for the hard-core attractive Yukawa system". *J. Chem. Phys.*, **101** (5), (1994) pp. 4093–4097.

[17] M. Noro and D. Frenkel. "Extended corresponding-states behavior for particles with variable range attractions". *J. Chem. Phys.*, **113**, (2000) p. 2941.

[18] G. Foffi and F. Sciortino. "On the possibility of extending the Noro-Frenkel generalized law of correspondent states to non-isotropic patchy interactions". *J. Phys. Chem. B*, **111**, (2007) p. 9702.

[19] K. N. Pham, A. M. Puertas, J. Bergenholtz, S. U. Egelhaaf, A. Moussad, P. N. Pusey, A. B. Schofield, M. E. Cates, M. Fuchs, and W. C. K. Poon. "Multiple Glassy States in a Simple Model System". *Science*, **296** (5565), (2002) pp. 104–106.

[20] F. Sciortino. "Disorderd materials: One liquid, two glasses". *Nat. Mater.*, **1** (3), (2002) pp. 145–146.

[21] M. Sperl, E. Zaccarelli, F. Sciortino, P. Kumar, and H. E. Stanley. "Disconnected Glass-Glass Transitions and Diffusion Anomalies in a Model with Two Repulsive Length Scales". *Phys. Rev. Lett.*, **104**, (2010) p. 145701.

[22] K. Dawson, G. Foffi, M. Fuchs, W. Götze, F. Sciortino, M. Sperl, P. Tartaglia, T. Voigtmann, and E. Zaccarelli. "Higher-order glass-transition singularities in colloidal systems with attractive interactions". *Phys. Rev. E*, **63**, (2000) p. 011401.

[23] Y.-S. Cho, G.-R.Yi, J.-M. Lim, S.-H. Kim, V. N. Manoharan, D. J. Pine, and S.-M. Yang. "Self-organization of bidisperse colloids in water droplets." *J. Am. Chem. Soc.*, **127**, (2005) pp. 15968–15975.

[24] Y. Sun and Y. Xia. "Shape-Controlled Synthesis of Gold and Silver Nanoparticles". *Science*, **298** (5601), (2002) pp. 2176–2179.

[25] L. Rossi, S. Sacanna, W. T. M. Irvine, P. M. Chaikin, D. J. Pine, and A. P. Philipse. "Cubic crystals from cubic colloids". *Soft Matter*, **7**, (2011) p. 4139.

[26] K.-H. Roh, D. C. Martin, and J. Lahann. "Biphasic Janus particles with nanoscale anisotropy". *Nat. Mater.*, **4**, (2005) pp. 759–763.

[27] C.-H. Chen, R. K. Shah, A. R. Abate, and D. A. Weitz. "Janus Particles Templated from Double Emulsion Droplets Generated Using Microfluidics". *Langmuir*, **25** (8), (2009) pp. 4320–4323.

[28] Q. Chen, S. Bae, and S. Granick. "Directed self-assembly of a colloidal kagome lattice". *Nature*, **469** (7330), (2011) pp. 381–384.

[29] V. N. Manoharan, M. T. Elsesser, and D. J. Pine. "Dense packing and symmetry in small clusters of microspheres." *Science*, **301**, (2003) pp. 483–487.

[30] G. Zhang, D. Wang, and H. Mhwald. "Patterning Microsphere Surfaces by Templating Colloidal Crystals". *Nano Lett.*, **5**, (2005) p. 143.

[31] M. Klokkenburg, C. Vonk, E. M. Claesson, J. D. Meeldijk, B. H. Ern, and A. P. Philipse. "Direct Imaging of Zero-Field Dipolar Structures in Colloidal Dispersions of Synthetic Magnetite". *JACS*, **126** (51), (2004) pp. 16706–16707.

[32] Z. Zhang and S. C. Glotzer. "Self-assembly patchy particles". *Nano Lett.*, **4** (8), (2004) pp. 1407–1413.

[33] F. Romano, E. Sanz, and F. Sciortino. "Role of the Range in the Fluid-Crystal Coexistence for a Patchy Particle Model". *J. Phys. Chem. B*, **113** (46), (2009) p. 15133.

[34] F. Romano, E. Sanz, and F. Sciortino. "Crystallization of tetrahedral patchy particles in silico". *J. Chem. Phys.*, **134** (17), 174502.

[35] F. Romano and F. Sciortino. "Patterning symmetry in the rational design of colloidal crystals". *Nat. Commun.*, **3**, (2012) p. 975.

[36] G. Doppelbauer, E. G. Noya, E. Bianchi, and G. Kahl. "Self-assembly scenarios of patchy colloidal particles". *Soft Matter*, **8**, (2012) pp. 7768–7772.

[37] F. Sciortino, E. Bianchi, J. F. Douglas, and P. Tartaglia. "Self-assembly of patchy particles into polymer chains: A parameter-free comparison between Wertheim theory and Monte Carlo simulation". *J. Chem. Phys.*, **126**, (2007) p. 194903.

[38] E. Zaccarelli, S. V. Buldyrev, E. L. Nave, A. J. Moreno, Saika-Voivod, F. Sciortino, and P. Tartaglia. "Model for reversible colloidal gelation." *Phys. Rev. Lett.*, **94**, (2005) p. 218301.

[39] E. Bianchi, J. Largo, P. Tartaglia, E. Zaccarelli, and F. Sciortino. "Phase diagram of patchy colloids: towards empty liquids". *Phys. Rev. Lett.*, **97**, (2006) pp. 168301–168304.

[40] L. Rovigatti, W. Kob, and F. Sciortino. "The vibrational density of states of a disordered gel model". *J. Chem. Phys.*, **135** (10), (2011) p. 104502.

[41] B. Ruzicka, E. Zaccarelli, L. Zulian, R. Angelini, M. Sztucki, A. Moussaïd, T. Narayanan, and F. Sciortino. "Observation of empty liquids and equilibrium gels in a colloidal clay". *Nat. Mater.*, **10** (1), (2010) pp. 56–60.

[42] F. Sciortino, S. V. Buldyrev, C. D. Michele, G. Foffi, N. Ghofraniha, E. L. Nave, A. Moreno, S. Mossa, I. Saika-Voivod, P. Tartaglia, and E. Zaccarelli. "Routes to Colloidal Gel Formation." *Comp. Phys. Comm.*, **169**, (2005) pp. 166–171.

[43] M. P. Allen. "Liquid Crystal Systems". In N. Attig, K. Binder, H. Grubmüller, and K. Kremer, editors, "Computational Soft Matter: From Synthetic Polymers to Proteins", volume 23, pp. 289–320. John von Neumann Institute for Computing, Jülich, **2004**.

[44] F. Reinitzer. "Beitrge zur Kenntniss des Cholesterins". *Monatshefte fr Chemie / Chemical Monthly*, **9**, (1888) pp. 421–441.

[45] L. Onsager. "THE EFFECTS OF SHAPE ON THE INTERACTION OF COLLOIDAL PARTICLES". *Annals of the New York Academy of Sciences*, **51**, (1949) p. 627.

[46] S.-D. Lee. "The Onsager-type theory for nematic ordering of finite-length hard ellipsoids". *J. Chem. Phys.*, **89**, (1989) p. 7036.

[47] A. Samborski, G. T. Evans, C. P. Mason, and M. P. Allen. "The isotropic to nematic liquid crystal transition for hard ellipsoids: An Onsager-like theory and computer simulations". *Mol. Phys.*, **81**, (1994) pp. 263–276.

[48] B. Tijpto-Margo and G. T. Evans. "The Onsager theory of the isotropic-nematic liquid crystal transition: incorporation of the higher virial coefficients". *J. Chem. Phys.*, **93**, (1990) p. 4254.

[49] M. Nakata, G. Zanchetta, B. D. Chapman, C. D. Jones, J. O. Cross, R. Pindak, T. Bellini, and N. A. Clark. "End-to-End Stacking and Liquid Crystal Condensation of 6 to 20 Base Pair DNA Duplexes". *Science*, **318**, (2007) p. 1276.

[50] J. SantaLucia and D. Hicks. "The thermodynamics of DNA structural motifs". *Annu. Rev. Biophys. Biomol. Struct.*, **33**, (2004) pp. 415–440.

[51] T. Bellini, R. Cerbino, and G. Zanchetta. "DNA-Based Soft Phases". In C. Tschierske, editor, "Liquid Crystals - Materials Design and Self-Assembly", volume 318 of *Topics in Current Chemistry*, pp. 225–279. Springer Berlin / Heidelberg, **2012**. To appear.

[52] C. De Michele, T. Bellini, and F. Sciortino. "Self-Assembly of Bifunctional Patchy Particles with Anisotropic Shape into Polymers Chains: Theory, Simulations, and Experiments". *Macromolecules*, **45** (2), (2012) pp. 1090–1106.

[53] C. De Michele, L. Rovigatti, T. Bellini, and F. Sciortino. "Self-assembly of short DNA duplexes: from a coarse-grained model to experiments through a theoretical link". *Soft Matter*, **8**, (2012) p. 8388.

[54] J. M. Tavares, L. Rovigatti, and F. Sciortino. "Quantitative description of the self-assembly of patchy particles into chains and rings". *J. Chem. Phys.*, **137** (4), (2012) p. 044901.

[55] T. Tlusty and S. A. Safran. "Defect-Induced Phase Separation in Dipolar Fluids". *Science*, **290**, (2000) p. 1328.

[56] J. Russo, J. M. Tavares, P. I. C. Teixeira, M. M. Telo da Gama, and F. Sciortino. "Reentrant Phase Diagram of Network Fluids". *Phys. Rev. Lett.*, **106**, (2011) p. 085703.

[57] J. Russo, J. M. Tavares, P. I. C. Teixeira, M. M. Telo da Gama, and F. Sciortino. "Re-entrant phase behaviour of network fluids: A patchy particle model with temperature-dependent valence". *J. Chem. Phys.*, **135**, (2011) p. 034501.

[58] P. G. de Gennes and P. A. Pincus. "Pair correlations in a ferromagnetic colloid". *Physik der Kondensierten Materie*, **11**, (1970) pp. 189–198.

[59] J. J. Weis and D. Levesque. "Chain formation in low density dipolar hard spheres: A Monte Carlo study". *Phys. Rev. Lett.*, **71**, (1993) pp. 2729–2732.

[60] P. I. C. Teixeira, J. M. Tavares, and M. M. Telo da Gama. "REVIEW ARTICLE: The effect of dipolar forces on the structure and thermodynamics of classical fluids". *J. Phys.: Condens. Matter*, **12**, (2000) p. 411.

[61] L. Rovigatti, J. Russo, and F. Sciortino. "No Evidence of Gas-Liquid Coexistence in Dipolar Hard Spheres". *Phys. Rev. Lett.*, **107**, (2011) p. 237801.

[62] L. Rovigatti, J. Russo, and F. Sciortino. "Structural properties of the dipolar hard-sphere fluid at low temperatures and densities". *Soft Matter*, **8**, (2012) p. 6310.

[63] F. Sciortino. "Gel-forming patchy colloids and network glass formers: thermodynamic and dynamic analogies". *Eur. Phys. J. B*, **64**, (2008) pp. 505–509.

[64] C. De Michele, S. Gabrielli, P. Tartaglia, and F. Sciortino. "Dynamics in the presence of attractive patchy interactions". *J. Phys. Chem. B*, **110**, (2006) p. 8064.

[65] L. Rovigatti and F. Sciortino. "Self and collective correlation functions in a gel of tetrahedral patchy particles". *Mol. Phys.*, **109** (23-24), (2011) p. 2889.

[66] G. M. Whitesides and M. Boncheva. "Beyond molecules: Self-assembly of mesoscopic and macroscopic components". *Proc. Natl. Acad. Sci. USA*, **99** (8), (2002) pp. 4769–4774.

[67] F. W. Starr and F. Sciortino. "LETTER TO THE EDITOR: Model for assembly and gelation of four-armed DNA dendrimers". *J. Phys.: Condens. Matter*, **18**, (2006) pp. L347–L353.

[68] Z. Nie, D. Fava, E. Kumacheva, S. Zou, G. C. Walker, and M. Rubinstein. "Self-assembly of metal-polymer analogues of amphiphilic triblock copolymers". *Nat. Mater.*, **6** (8), (2007) pp. 609–614.

[69] J. P. K. Doye, A. A. Louis, I.-C. Lin, L. R. Allen, E. G. Noya, A. W. Wilber, H. C. Kok, and R. Lyus. "Controlling crystallization and its absence: Proteins, colloids and patchy models". *Phys. Chem. Chem. Phys.*, **9**, (2007) pp. 2197–2205.

[70] C. Mirkin, R. Letsinger, R. Mucic, and J. Storhoffand. "A DNA-based method for rationally assembling nanoparticles into macroscopic materials". *Nature*, **382**, (1996) pp. 607–609.

[71] V. Workum and J. Douglas. "Symmetry, equivalence, and molecular self-assembly". *Phys. Rev. E*, **73**, (2006) p. 031502.

[72] A. Khan. "Phase science of surfactants". *Current Opinion in Colloid & Interface Science*, **1**, (1996) p. 614.

[73] P. van der Schoot and M. Cates. "Growth, Static Light Scattering, and Spontaneous Ordering of Rodlike Micelles". *Langmuir*, **10**, (1994) pp. 670–679.

[74] D. M. Kuntz and L. M. Walker. "Nematic phases observed in amphiphilic polyelectrolyte-surfactant aggregate solutions". *Soft Matter*, **4**, (2008) pp. 286–293.

[75] J.-M. Jung and R. Mezzenga. "Liquid Crystalline Phase Behavior of Protein Fibers in Water: Experiments versus Theory". *Langmuir*, **26** (1), (2010) pp. 504–514.

[76] C. F. Lee. "Isotropic-nematic phase transition in amyloid fibrilization". *Phys. Rev. E*, **80** (3), (2009) p. 031902.

[77] A. Ciferri. "On collagen II fibrillogenesis". *Liq. Cryst.*, **34** (6), (2007) pp. 693–696.

[78] A. Aggeli, M. Bell, L. M. Carrick, C. W. G. Fishwick, R. Harding, P. J. Mawer, S. E. Radford, A. E. Strong, and N. Boden. "pH as a Trigger of Peptide -Sheet Self-Assembly and Reversible Switching between Nematic and Isotropic Phases". *J. Am. Chem. Soc.*, **125** (32), (2003) pp. 9619–9628.

[79] C. Robinson. "Liquid-crystalline structures in polypeptide solutions". *Tetrahedron*, **13** (1-3), (1961) pp. 219 – 234.

[80] F. Livolant, A. M. Levelut, J. Doucet, and J. P. Benoit. "The highly concentrated liquid-crystalline phase of DNA is columnar hexagonal". *Nature*, **339** (6227), (1989) pp. 724–726.

[81] K. Merchant and R. L. Rill. "DNA length and concentration dependencies of anisotropic phase transitions of DNA solutions". *Biophys. J.*, **73** (6), (1997) pp. 3154–3163.

[82] F. Tombolato and A. Ferrarini. "From the double-stranded helix to the chiral nematic phase of B-DNA: A molecular model". *J. Chem. Phys.*, **122** (5), (2005) p. 054908.

[83] F. Tombolato, A. Ferrarini, and E. Grelet. "Chiral Nematic Phase of Suspensions of Rodlike Viruses: Left-Handed Phase Helicity from a Right-Handed Molecular Helix". *Phys. Rev. Lett.*, **96** (25), (2006) p. 258302.

[84] E. Barry, D. Beller, and Z. Dogic. "A model liquid crystalline system based on rodlike viruses with variable chirality and persistence length". *Soft Matter*, **5**, (2009) pp. 2563–2570.

[85] E. Grelet and S. Fraden. "What Is the Origin of Chirality in the Cholesteric Phase of Virus Suspensions?" *Phys. Rev. Lett.*, **90** (19), (2003) p. 198302.

[86] S. Tomar, M. M. Green, and L. A. Day. "DNAàíProtein Interactions as the Source of Large-Length-Scale Chirality Evident in the Liquid Crystal Behavior of Filamentous Bacteriophages". *J. Am. Chem. Soc.*, **129** (11), (2007) pp. 3367–3375.

[87] A. Minsky, E. Shimoni, and D. Frenkiel-Krispin. "Stress, order and survival". *Nat. Rev. Mol. Cell. Biol.*, **3** (1), (2002) pp. 50–60.

[88] J. Lydon. "Chromonic review". *J. Mater. Chem.*, **20**, (2010) pp. 10071–10099.

[89] K. Liu, Z. Nie, N. Zhao, W. Li, M. Rubinstein, and E. Kumacheva. "Step-Growth Polymerization of Inorganic Nanoparticles". *Science*, **329** (5988), (2010) pp. 197–200.

[90] G. Zanchetta, M. Nakata, M. Buscaglia, N. A. Clark, and T. Bellini. "Liquid crystal ordering of DNA and RNA oligomers with partially overlapping sequences". *J. Phys.: Condens. Matter*, **20** (49), (2008) p. 494214.

[91] G. Zanchetta, F. Giavazzi, M. Nakata, M. Buscaglia, R. Cerbino, N. A. Clark, and T. Bellini. "Right-handed double-helix ultrashort DNA yields chiral nematic phases with both right- and left-handed director twist". *Proc. Natl. Acad. Sci. USA*, **107** (41), (2010) pp. 17497–17502.

[92] K. M. Guckian, B. A. Schweitzer, R. X.-F. Ren, C. J. Sheils, D. C. Tahmassebi, and E. T. Kool. "Factors Contributing to Aromatic Stacking in Water:Äâ Evaluation in the Context of DNA". *J. Am. Chem. Soc.*, **122** (10), (2000) pp. 2213–2222.

[93] T. Kuriabova, M. Betterton, and M. Glaser. "Linear aggregation and liquid-crystalline order: comparison of Monte Carlo simulation and analytic theory". *J. Mater. Chem.*, **20**, (2010) pp. 10366–10383.

[94] X. Lü and J. Kindt. "Monte Carlo simulation of the self-assembly and phase behavior of semiflexible equilibrium polymers". *J. Chem. Phys.*, **120**, (2004) pp. 10328–10338.

[95] G. J. Vroege and H. N. W. Lekkerkerker. "Phase transitions in lyotropic colloidal and polymer liquid crystals". *Rep. Prog. Phys.*, **55**, (1992) pp. 1241–1309.

[96] M. Dijkstra and D. Frenkel. "Simulation study of the isotropic-to-nematic transitions of semiflexible polymers". *Phys. Rev. E*, **51** (6), (1995) pp. 5891–5898.

[97] A. Khokhlov and A. Semenov. "Liquid-Crystalline ordering in the solution of long persistent chains". *Physica*, **108A**, (1981) pp. 546–556.

[98] A. Khokhlov and A. Semenov. "Liquid-crystalline ordering in the solution of partially flexibly macromolecules". *Physica*, **112A**, (1982) pp. 605–614.

[99] P. P. F. Wessels and B. M. Mulder. "Isotropic-to-nematic transition in liquid-crystalline heteropolymers: I. Formalism and main-chain liquid-crystalline polymers". *J. Phys.: Condens. Matter*, **18** (41), (2006) p. 9335.

[100] M. Dennison, M. Dijkstra, and R. van Roij. "Phase Diagram and Effective Shape of Semiflexible Colloidal Rods and Biopolymers". *Phys. Rev. Lett.*, **106** (20), (2011) p. 208302.

[101] Z. Wang, D. Kuckling, and D. Johannsmann. "Temperature-Induced Swelling and De-swelling of Thin Poly(N-Isopropylacrylamide) Gels in

Water: Combined Acoustic and Optical Measurements". *Soft Mater.*, **1** (3), (2003) pp. 353–364.

[102] Z. Y. Chen. "Nematic ordering in semiflexible polymer chains". *Macromolecules*, **26** (13), (1993) pp. 3419–3423.

[103] T. Odijk. "Theory of Lyotropic Polymer Liquid Crystals". *Macromolecules*, **19**, (1986) p. 2313.

[104] C. De Michele. "Simulating Hard Rigid Bodies". *J. Comput. Phys.*, **229**, (2010) pp. 3276–3294.

[105] F. Sciortino, C. De Michele, S. Corezzi, J. Russo, E. Zaccarelli, and P. Tartaglia. "A parameter-free description of the kinetics of formation of loop-less branched structures and gels". *Soft Matter*, **5**, (2009) pp. 2571–2575.

[106] S. Corezzi, C. De Michele, E. Zaccarelli, P. Tartaglia, and F. Sciortino. "Connecting Irreversible to Reversible Aggregation: Time and Temperature". *J. Phys. Chem. B*, **113** (5), (2009) pp. 1233–1236.

[107] M. Wertheim. "Fluids with Highly Directional Attractive Forces. I. Statistical Thermodynamics". *J. Stat. Phys.*, **35**, (1984) pp. 19–34.

[108] M. Wertheim. "Fluids with Highly Directional Attractive Forces. II. Thermodynamic Perturbation Theory and Integral Equations". *J. Stat. Phys.*, **35**, (1984) pp. 35–47.

[109] M. Wertheim. "Fluids with Highly Directional Attractive Forces. III. Multiple Attraction Sites". *J. Stat. Phys.*, **42**, (1986) pp. 459–476.

[110] T. E. Ouldridge, A. A. Louis, and J. P. K. Doye. "Structural, mechanical, and thermodynamic properties of a coarse-grained DNA model". *J. Chem. Phys.*, **134** (8), (2011) p. 085101.

[111] T. E. Ouldridge, A. A. Louis, and J. P. K. Doye. "DNA Nanotweezers Studied with a Coarse-Grained Model of DNA". *Phys. Rev. Lett.*, **104**, (2010) p. 178101.

[112] T. E. Ouldridge. *Coarse-Grained Modelling of DNA and DNA self-assembly*. Ph.D. thesis, University of Oxford (**2011**).

[113] E. Protozanova, P. Yakovchuk, and M. Frank-Kamenetskii. "Stacked-unstacked equilibrium at the nick site of DNA." *J. Mol. Biol.*, **342** (3), (2004) pp. 775–785.

[114] P. Yakovchuk, E. Protozanova, and M. D. Frank-Kamenetskii. "Base-stacking and base-pairing contributions into thermal stability of the DNA double helix". *Nucl. Acids Res.*, **34** (2), (2006) pp. 564–574.

[115] S. Bommarito, N. Peyret, and J. J. SantaLucia. "Thermodynamic parameters for DNA sequences with dangling ends". *Nucl. Acids Res.*, **28** (9), (2000) pp. 1929–1934.

[116] J. SantaLucia. "A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics". *Proc. Natl. Acad. Sci. USA*, **95** (4), (1998) p. 1460.

[117] P. Šulc, F. Romano, T. E. Ouldridge, L. Rovigatti, A. A. Louis, and J. P. K. Doye. "Introducing sequence-dependent interactions into a coarse-grained DNA model". *J. Chem. Phys.*, **137** (13), (2012) p. 135101.

[118] http://dna.physics.ox.ac.uk.

[119] J. Parsons. "Nematic ordering in a system of rods". *Phys. Rev. A*, **19**, (1979) pp. 1225–1230.

[120] C. Maffeo, B. Luan, and A. Aksimentiev. "End-to-end attraction of duplex DNA". *Nucleic Acids Research*, **40** (9), (2012) pp. 3812–3821.

[121] E. Frezza, F. Tombolato, and A. Ferrarini. "Right- and left-handed liquid crystal assemblies of oligonucleotides: phase chirality as a reporter of a change in non-chiral interactions?" *Soft Matter*, **7**, (2011) pp. 9291–9296.

[122] M. E. Cates and S. J. Candau. "REVIEW ARTICLE: Statics and dynamics of worm-like surfactant micelles". *J. Phys.: Condens. Matter*, **2**, (1990) pp. 6869–6892.

[123] A. Milchev and Y. Rouault. "A Monte Carlo Study of Thermodynamic Relaxation in Living Polymers". *Journal de Physique II*, **5**, (1995) pp. 343–347.

[124] A. N. Semenov, I. A. Nyrkova, and M. E. Cates. "Phase equilibria in solutions of associating telechelic polymers: rings vs reversible network". *Macromolecules*, **28** (23), (1995) pp. 7879–7885.

[125] A. Milchev, J. P. Wittmer, and D. P. Landau. "Dynamical Monte Carlo study of equilibrium polymers: Effects of high density and ring formation". *Phys. Rev. E*, **61**, (2000) pp. 2959–2966.

[126] J. P. Wittmer, P. van der Schoot, A. Milchev, and J. L. Barrat. "Dynamical Monte Carlo study of equilibrium polymers. II. The role of rings". *J. Chem. Phys.*, **113** (16), (2000) pp. 6992–7005.

[127] J. A. Cuesta and R. P. Sear. "Phase transition analogous to Bose-Einstein condensation in systems of noninteracting surfactant aggregates". *Phys. Rev. E*, **65** (3), (2002) p. 031406.

[128] R. P. Sear and G. Jackson. "Thermodynamic perturbation theory for association into chains and rings". *Phys. Rev. E*, **50**, (1994) pp. 386–394.

[129] M. Rubinstein and R. H. Colby. *Polymer Physics*. Oxford University Press Inc., New York, **2003**.

[130] M. Wertheim. "Fluids with highly directional attractive forces. I. Statistical thermodynamics". *J. Stat. Phys.*, **35**, (1984) pp. 19, ibid. 35.

[131] M. Wertheim. *J. Stat. Phys.*, **42**, (1986) pp. 459, ibid. 477.

[132] M. S. Wertheim. "Thermodynamic perturbation theory of polymerization". *J. Chem. Phys.*, **87**, (1987) p. 7323.

[133] F. Sciortino, E. Bianchi, J. F. Douglas, and P. Tartaglia. "Self-assembly of patchy particles into polymer chains: A parameter-free comparison between Wertheim theory and Monte Carlo simulation". *J. Chem. Phys.*, **126**, (2007) p. 4903.

[134] A. Galindo, S. Burton, G. Jackson, D. Visco, and D. A. Kofke. *Mol. Phys.*, **100**, (2002) p. 2241.

[135] A. Avlund, G. Kontogeorgis, and W. Chapman. *Mol. Phys.*, **109**, (2011) p. 1759.

[136] S. Kantorovich, J. Cerda, and C. Holm. "Microstructure analysis of monodisperse ferrofluid monolayers: theory and simulation". *Phys. Chem. Chem. Phys.*, **10** (14), (2008) pp. 1883–1895.

[137] D. Tománek, S. Kim, P. Jund, P. Borrmann, H. Stamerjohanns, and E. Hilf. "Self-assembly of magnetic nanostructures". *ZEITSCHRIFT FUR PHYSIK D ATOMS MOLECULES AND CLUSTERS*, **40**, (1997) pp. 539–541.

[138] F. Kun, W. Wen, K. Pál, and K. Tu. "Breakup of dipolar rings under a perpendicular magnetic field". *Physical Review E*, **64** (6), (2001) p. 061503.

[139] H. Morimoto, T. Maekawa, and Y. Matsumoto. "Statistical analysis of two-dimensional cluster structures composed of ferromagnetic particles based on a flexible chain model". *Phys. Rev. E*, **68**, (2003) p. 061505.

[140] X. Chen, M. Yu-Qiang, H. Pak-Ming, and T. Fu-Qiang. "Microstructures in strongly interacting dipolar fluids". *Chinese Physics Letters*, **22**, (2005) p. 485.

[141] Y. Koksharov, G. Khomutov, E. Soldatov, D. Suyatin, I. Maximov, L. Montelius, and P. Carlberg. "Magnetostatic interactions in planar ring-like nanoparticle structures". *Thin solid films*, **515** (2), (2006) pp. 731–734.

[142] J. Tavares, P. Teixeira, and M. T. da Gama. *Mol. Phys.*, **107**, (2009) p. 453.

[143] J. Tavares and P. Teixeira. "Patching up dipoles: can dipolar particles be viewed as patchy colloids?" *Mol. Phys.*, **109** (7-10), (2011) pp. 1077–1085.

[144] N. Kern and D. Frenkel. "Fluid-fluid coexistence in colloidal systems with short-ranged strongly directional attraction." *J. Chem. Phys.*, **118**, (2003) pp. 9882–9889.

[145] E. Zaccarelli, I. Saika-Voivod, S. V. Buldyrev, A. J. Moreno, P. Tartaglia, and F. Sciortino. "Gel to glass transition in simulation of a valence-limited colloidal system." *J. Chem. Phys.*, **124**, (2006) p. 124908.

[146] B. Chen and J. I. Siepmann. "Improving the Efficiency of the Aggregation-Volume-Bias Monte Carlo Algorithm". *J. Phys. Chem. B*, **105** (45), (2001) pp. 11275–11282.

[147] P. J. Flory. *Statistical Mechanics of Chain Molecules*. Interscience publishers (New York and London), **1969**.

[148] G. Porte. *Micelles, Membranes, Microemulsions, and Monolayers, Eds. W. M. Gelbart, A. Ben-Shaul, and D. Roux*. Springer-Verlag, New York, **1994**.

[149] J. M. Tavares, J. J. Weis, and M. M. Telo da Gama. "Quasi-two-dimensional dipolar fluid at low densities: Monte Carlo simulations and theory". *Phys. Rev. E*, **65** (6), (2002) p. 061201.

[150] J.-M. Caillol. "Search of the gas–liquid transition of dipolar hard spheres". *J. Chem. Phys.*, **98** (12), (1993) pp. 9835–9849.

[151] G. Ganzenmüller, G. N. Patey, and P. J. Camp. "Vapour-liquid phase transition of dipolar particles". *Mol. Phys.*, **107**, (2009) pp. 403–413.

[152] M. E. van Leeuwen and B. Smit. *Phys. Rev. Lett.*, **71** (24), (1993) p. 3991.

[153] J.-J. Weis and D. Levesque. "Simple Dipolar Fluids as Generic Models for Soft Matter". In C. Holm and K. Kremer, editors, "Advanced Computer Simulation Approaches for Soft Matter Sciences II", volume 185 of *Advances in Polymer Science*, pp. 163–225. Springer Berlin Heidelberg, **2005**.

[154] R. van Roij. "Theory of Chain Association versus Liquid Condensation". *Phys. Rev. Lett.*, **76** (18), (1996) pp. 3348–3351.

[155] R. P. Sear. "Low-Density Fluid Phase of Dipolar Hard Spheres". *Phys. Rev. Lett.*, **76**, (1996) p. 2310.

[156] J. Dudowicz, K. F. Freed, and J. F. Douglas. *J. Chem. Phys.*, **119**, (2003) p. 12645.

[157] J. Dudowicz, K. F. Freed, and J. F. Douglas. "Flory-Huggins model of equilibrium polymerization and phase separation in the Stockmayer fluid". *Phys. Rev. Lett.*, **92**, (2004) pp. 045502–045505.

[158] K. Ng, J. P. Valleau, G. M. Torrie, and G. N. Patey. "Liquid-vapour coexistence of dipolar hard spheres". *Mol. Phys.*, **38**, (1979) pp. 781–788.

[159] Kalyuzhnyi, Y. V., Protsykevytch, I. A., and Cummings, P. T. "Thermodynamic properties and liquid-gas phase diagram of the dipolar hard-sphere fluid". *EPL*, **80** (5), (2007) p. 56002.

[160] J. C. Shelley, G. N. Patey, D. Levesque, and J. J. Weis. "Liquid-vapor coexistence in fluids of dipolar hard dumbbells and spherocylinders". *Phys. Rev. E*, **59** (3), (1999) pp. 3065–3070.

[161] N. G. Almarza, E. Lomba, C. Martín, and A. Gallardo. "Demixing in binary mixtures of apolar and dipolar hard spheres". *J. Chem. Phys.*, **129** (23), (2008) p. 234504.

[162] G. Ganzenmüller and P. J. Camp. "Vapor-liquid coexistence in fluids of charged hard dumbbells". *J. Chem. Phys.*, **126**, (2007) p. 191104.

[163] M. Klokkenburg, B. Erné, A. Wiedenmann, A. Petukhov, and A. Philipse. "Dipolar structures in magnetite ferrofluids studied with small-angle neutron scattering with and without applied magnetic field". *Phys. Rev. E*, **75** (5), (2007) p. 051408.

[164] A. Wiedenmann, U. Keiderling, M. Meissner, D. Wallacher, R. Gähler, R. P. May, S. Prévost, M. Klokkenburg, B. H. Erné, and J. Kohlbrecher. "Low-temperature dynamics of magnetic colloids studied by time-resolved small-angle neutron scattering". *Phys. Rev. B*, **77**, (2008) p. 184417.

[165] M. Klokkenburg, R. P. A. Dullens, W. K. Kegel, B. H. Erné, and A. P. Philipse. "Quantitative Real-Space Analysis of Self-Assembled Structures of Magnetic Dipolar Colloids". *Phys. Rev. Lett.*, **96** (3), (2006) p. 037203.

[166] A. O. Ivanov, Z. Wang, and C. Holm. "Applying the chain formation model to magnetic properties of aggregated ferrofluids". *Phys. Rev. E*, **69**, (2004) p. 031206.

[167] J. Cerdà, E. Elfimova, V. Ballenegger, E. Krutikova, A. Ivanov, and C. Holm. "Behavior of bulky ferrofluids in the diluted low-coupling regime: Theory and simulation". *Physical Review E*, **81** (1), (2010) p. 011501.

[168] C. Holm, A. Ivanov, S. Kantorovich, E. Pyanzina, and E. Reznikov. "Equilibrium properties of a bidisperse ferrofluid with chain aggregates: theory and computer simulations". *Journal of Physics: Condensed Matter*, **18** (38), (2006) p. S2737.

[169] E. Pyanzina, S. Kantorovich, J. Cerdà, A. Ivanov, and C. Holm. "How to analyse the structure factor in ferrofluids with strong magnetic interactions: a combined analytic and simulation approach". *Molecular Physics*, **107** (4-6), (2009) pp. 571–590.

[170] P. J. Camp and G. N. Patey. "Structure and scattering in colloidal ferrofluids". *Phys. Rev. E*, **62**, (2000) pp. 5403–5408.

[171] T. Prokopieva, V. Danilov, S. Kantorovich, and C. Holm. "Ground state structures in ferrofluid monolayers". *Physical Review E*, **80** (3), (2009) p. 031404.

[172] B. Smith and D. Frenkel. *Understanding molecular simulations*. Academic Press, New York, **1996**.

[173] P. Virnau and M. Müller. "Calculation of free energy through successive umbrella sampling". *J. Chem. Phys.*, **120** (23), (2004) pp. 10925–10930.

[174] B. J. Schulz, K. Binder, M. Müller, and D. P. Landau. "Avoiding boundary effects in Wang-Landau sampling". *Phys. Rev. E*, **67** (6), (2003) p. 067102.

[175] A. M. Ferrenberg and R. H. Swendsen. *Phys. Rev. Lett.*, **61**, (1988) p. 2635.

[176] D. Levesque and J. J. Weis. "Orientational and structural order in strongly interacting dipolar hard spheres". *Phys. Rev. E*, **49** (6), (1994) pp. 5131–5140.

[177] J. M. Tavares, J. J. Weis, and M. M. Telo da Gama. "Strongly dipolar fluids at low densities compared to living polymers". *Phys. Rev. E*, **59** (4), (1999) pp. 4388–4395.

[178] Z. Wang, C. Holm, and H. W. Müller. "Boundary condition effects in the simulation study of equilibrium properties of magnetic dipolar fluids". *J. Chem. Phys.*, **119** (1), (2003) pp. 379–387.

[179] R. Jia, H. Braun, and R. Hentschke. "Gas-liquid coexistence in a system of dipolar soft spheres". *Phys. Rev. E*, **82** (6), (2010) p. 062501.

[180] M. A. Floriano, E. Caponetti, and A. Z. Panagiotopoulos. "Micellization in Model Surfactant Systems". *Langmuir*, **15** (9), (1999) pp. 3143–3151.

[181] F. Lo Verso, A. Z. Panagiotopoulos, and C. N. Likos. "Aggregation phenomena in telechelic star polymer solutions". *Phys. Rev. E*, **79** (1), (2009) p. 010401.

[182] F. Sciortino, A. Giacometti, and G. Pastore. "Phase Diagram of Janus Particles". *Phys. Rev. Lett.*, **103**, (2009) p. 237801.

[183] A. Reinhardt, A. J. Williamson, J. P. K. Doye, J. Carrete, L. M. Varela, and A. A. Louis. "Re-entrant phase behavior for systems with competition between phase separation and self-assembly". *J. Chem. Phys.*, **134**, (2011) p. 104905.

[184] R. Jones. *Soft Condensed Matter*. Oxford University Press, **2002**.

[185] F. H. Stillinger and T. A. Weber. "Inherent structure in water". *The Journal of Physical Chemistry*, **87** (15), (1983) pp. 2833–2840.

[186] F. Sciortino. "Potential energy landscape description of supercooled liquids and glasses". *Journal of Statistical Mechanics: Theory and Experiment*, **5**, (2005) p. 15.

[187] A. Coniglio and W. Klein. *J. Phys. A*, **13**, (1980) p. 2775.

[188] S. H. Chen, J. Rouch, F. Sciortino, and P. Tartaglia. "Static and dynamic properties of water-in-oil microemulsions near the critical and percolation points". *Journal of Physics: Condensed Matter*, **6** (50), (1994) p. 10855.

[189] E. Bianchi, P. Tartaglia, E. La Nave, and F. Sciortino. "Fully Solvable Equilibrium Self-Assembly Process: Fine-Tuning the Clusters Size and the Connectivity in Patchy Particle Systems". *J. Phys. Chem. B*, **111**, (2007) pp. 11765–11769.

[190] V. Workum and J. F. Douglas. *Phys. Rev. E*, **71**, (2005) p. 031502.

[191] P.-g. De Gennes. *Scaling Concepts in Polymer Physics*. Cornell University Press, **1980**.

[192] V. S. Mendelev and A. O. Ivanov. "Ferrofluid aggregation in chains under the influence of a magnetic field". *Phys. Rev. E*, **70**, (2004) p. 051502.

[193] J. P. Wittmer, A. Milchev, and M. E. Cates. "Dynamical Monte Carlo study of equilibrium polymers: Static properties". *J. Chem. Phys.*, **109**, (1998) pp. 834–845.

[194] Y. Rouault and A. Milchev. "Monte Carlo study of living polymers with the bond-fluctuation method". *Phys. Rev. E*, **51**, (1995) pp. 5905–5910.

[195] A. B. Pawar and I. Kretzschmar. "Fabrication, Assembly, and Application of Patchy Particles". *Macromol. Rapid Comm.*, **31**, (2010) pp. 150–168.

[196] E. Bianchi, R. Blaak, and C. N. Likos. "Patchy colloids: state of the art and perspectives". *Phys. Chem. Chem. Phys.*, **13**, (2011) pp. 6397–6410.

[197] S. Saw, N. L. Ellegaard, W. Kob, and S. Sastry. "Computer simulation study of the phase behavior and structural relaxation in a gel-former modeled by three-body interactions". *J. Chem. Phys.*, **134**, (2011) p. 164506.

[198] R. Blaak, M. A. Miller, and J.-P. Hansen. "Reversible gelation and dynamical arrest of dipolar colloids". *Europhys. Lett.*, **78**, (2007) p. 26002.

[199] E. Zaccarelli. "Colloidal Gels: Equilibrium and Non-Equilibrium Routes". *J. Phys.: Condens. Matter*, **19**, (2007) p. 323101.

[200] E. Del Gado and W. Kob. "Structure and relaxation dynamics of a colloidal gel". *Europhys. Lett.*, **72**, (2005) pp. 1032–1038.

[201] E. Del Gado and W. Kob. "Length-Scale-Dependent Relaxation in Colloidal Gels". *Phys. Rev. Lett.*, **98**, (2007) p. 028303.

[202] J. Horbach and W. Kob. "Static and dynamic properties of a viscous silica melt". *Phys. Rev. B*, **60**, (1999) pp. 3169–3181.

[203] F. W. Starr, F. Sciortino, and H. E. Stanley. "Dynamics of simulated water under pressure". *Phys. Rev. E*, **60**, (1999) pp. 6757–6768.

[204] http://www.nvidia.com/object/cuda_home_new.html.

[205] B. Ruzicka, E. Zaccarelli, L. Zulian, R. Angelini, M. Sztucki, A. Moussaïd, T. Narayanan, and F. Sciortino. "Observation of empty liquids and equilibrium gels in a colloidal clay". *Nat. Mater.*, **10**, (2011) pp. 56–60.

[206] C. A. Angell. "Relaxation in liquids, polymers and plastic crystals – strong/fragile patterns and problems". *J. Non-Cryst. Solids*, **131-133**, (1991) pp. 13 – 31.

[207] P. H. Poole, S. R. Becker, F. Sciortino, and F. W. Starr. "Dynamical Behavior Near a LiquidLiquid Phase Transition in Simulations of Supercooled Water". *The Journal of Physical Chemistry B*, **115** (48), (2011) p. 14176.

[208] W. Götze. *Liquids, Freezing and the Glass Transition*, pp. 287–503. North-Holland, Amsterdam, **1991**.

[209] F. Sciortino, L. Fabbian, S. H. Chen, and P. Tartaglia. "Supercooled water and the Kinetic glass transition. II Collective dynamics". *Phys. Rev. E*, **56**, (1997) pp. 5397–5404.

[210] A. Rinaldi, F. Sciortino, and P. Tartaglia. "Dynamics in a supercooled molecular liquid: Theory and simulations". *Phys. Rev. E*, **63**, (2001) p. 061210.

[211] J. Horbach and W. Kob. "Relaxation dynamics of a viscous silica melt: The intermediate scattering functions". *Phys. Rev. E*, **64**, (2001) p. 041503.

[212] F. Sciortino and W. Kob. "Debye-Waller Factor of Liquid Silica: Theory and Simulation". *Phys. Rev. Lett.*, **86**, (2001) pp. 648–651.

[213] E. Winfree, F. Liu, L. Wenzler, N. Seeman *et al.* "Design and self-assembly of two-dimensional DNA crystals". *Nature*, **394** (6693), (1998) pp. 539–544.

[214] N. Seeman. "DNA in a material world". *Nature*, **421** (6921), (2003) pp. 427–431.

[215] N. C. Seeman. "DNA Nanotechnology: Novel DNA Constructions". *Annual Review of Biophysics and Biomolecular Structure*, **27** (1), (1998) pp. 225–248.

[216] P. Rothemund. "Folding DNA to create nanoscale shapes and patterns". *Nature*, **440** (7082), (2006) pp. 297–302.

[217] N. Geerts and E. Eiser. "DNA-functionalized colloids: Physical properties and applications". *Soft Matter*, **6**, (2010) pp. 4647–4660.

[218] J. Largo, F. W. Starr, and F. Sciortino. *Langmuir*, **23**, (2007) pp. 5896–5905.

[219] S. Park, A. Lytton-Jean, B. Lee, S. Weigand, G. Schatz, and C. Mirkin. "DNA-programmable nanoparticle crystallization". *Nature*, **451** (7178), (2008) pp. 553–556.

[220] D. Nykypanchuk, M. Maye, D. Van Der Lelie, and O. Gang. "DNA-guided crystallization of colloidal nanoparticles". *Nature*, **451** (7178), (2008) pp. 549–552.

[221] B. M. Mladek, J. Fornleitner, F. J. Martinez-Veracoechea, A. Dawid, and D. Frenkel. "Quantitative Prediction of the Phase Diagram of DNA-Functionalized Nanosized Colloids". *Phys. Rev. Lett.*, **108**, (2012) p. 268301.

[222] D. Zanchet, C. M. Micheel, W. J. Parak, D. Gerion, and A. P. Alivisatos. "Electrophoretic Isolation of Discrete Au Nanocrystal/DNA Conjugates". *Nano Letters*, **1** (1), (2001) pp. 32–35.

[223] S. A. Claridge, H. W. Liang, S. R. Basu, J. M. J. Frchet, and A. P. Alivisatos. "Isolation of Discrete NanoparticleDNA Conjugates for Plasmonic Applications". *Nano Letters*, **8** (4), (2008) pp. 1202–1206. PMID: 18331002.

[224] K. Suzuki, K. Hosokawa, and M. Maeda. "Controlling the Number and Positions of Oligonucleotides on Gold Nanoparticle Surfaces". *Journal*

of the American Chemical Society, **131** (22), (2009) pp. 7518–7519. PMID: 19445511.

[225] S. Biffi and T. Bellini. Private communication.

[226] J. A. Holbrook, M. W. Capp, R. M. Saecker, and M. T. Record. "Enthalpy and Heat Capacity Changes for Formation of an Oligomeric DNA Duplex: Interpretation in Terms of Coupled Processes of Formation and Association of Single-Stranded Helices". *Biochemistry*, **38** (26), (1999) pp. 8409–8422.

[227] E. Sambriski, D. Schwartz, and J. de Pablo. "A Mesoscale Model of DNA and Its Renaturation". *Biophysical Journal*, **96** (5), (2009) pp. 1675 – 1690.

[228] T. A. Knotts, N. Rathore, D. C. Schwartz, and J. J. de Pablo. "A coarse grain model for DNA". *J. Chem. Phys.*, **126**, (2007) p. 4901.

[229] C. Schildkraut and S. Lifson. "Dependence of the melting temperature of DNA on salt concentration". *Biopolymers*, **3** (2), (1965) pp. 195–208.

[230] T. E. Ouldridge, A. A. Louis, and J. P. K. Doye. "Extracting bulk properties of self-assembling systems from small simulations". *J. Phys.: Cond. Matt.*, **22** (10), (2010) p. 104102.

[231] NVIDIA. *NVIDIA CUDA Programming Guide 4.2*. **2012**.

[232] T. Preis, P. Virnau, W. Paul, and J. J. Schneider. "GPU accelerated Monte Carlo simulation of the 2D and 3D Ising model". *Journal of Computational Physics*, **228** (12), (2009) pp. 4468 – 4477.

[233] M. Weigel. "Simulating spin models on GPU". *Computer Physics Communications*, **182** (9), (2011) p. 1833.

[234] J. A. Anderson, C. D. Lorenz, and A. Travesset. "General purpose molecular dynamics simulations fully implemented on graphics processing units". *J. Comput. Phys.*, **227**, (2008) pp. 5342–5359.

[235] M. P. Allen and D. J. Tildesley. *Computer simulation of liquids*. Clarendon Press, **1989**.

[236] P. H. Colberg and F. Höfling. "Highly accelerated simulations of glassy dynamics using GPUs: Caveats on limited floating-point precision". *Computer Physics Communications*, **182**, (2011) pp. 1120 – 1129.