INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **18** (2006) L347–L353 doi:10.1088/0953-8984/18/26/L02

LETTER TO THE EDITOR

Model for assembly and gelation of four-armed DNA dendrimers

Francis W Starr[1] **and Francesco Sciortino**[2]

[1] Department of Physics, Wesleyan University, Middletown, CT 06459, USA
[2] Dipartimento di Fisica and INFM Udr and Crs:Soft: Complex Dynamics in Structured Systems, Universitá di Roma La Sapienza, Piazzale Aldo Moro 2, I-00185 Rome, Italy

E-mail: fstarr@wesleyan.edu and francesco.sciortino@phys.uniroma1.it

Received 6 April 2006
Published 19 June 2006
Online at stacks.iop.org/JPhysCM/18/L347

Abstract

We introduce and numerically study a model designed to mimic the bulk behaviour of a system composed of single-stranded DNA dendrimers. Complementarity of the base sequences of different strands results in the formation of strong cooperative intermolecular links. We find that in an extremely narrow temperature range the system forms a large-scale, low-density disordered network via a thermo-reversible gel transition. By controlling the strand length, the gel transition temperature can be made arbitrarily close to the percolation transition, in contrast with recent model systems of physical gelation. This study helps the understanding of self-assembly in this class of new biomaterials and provides a bridge between physical and chemical gels.

(Some figures in this article are in colour only in the electronic version)

The 'bottom-up' construction of new materials is one of the central aims of nanotechnology [1]. The recent synthesis of nanoparticle building blocks functionalized with specifically designed oligonucleotides has opened new possibilities for the assembly of networked materials [2–5]. In this approach, a number of single strands of DNA are grafted on the surface of micron-sized particles. The interactions between the particles are controlled by the addition of DNA strands in solution which have complementary sequences to the DNA strands grafted on the particles. The novelty of the approach arises from the possibility of using the high sensitivity and selectivity of complementary strand recognition to turn on and off the interparticle bonding. The use of DNA sequences to establish interparticle interactions provides an optimal choice for the construction of three-dimensional supramolecular assemblies, since DNA strands can self-assemble into long and fairly rigid helices based on sequence complementarity [6, 7]. New materials can be designed by modulating the length of the binding sequence, the length of the grafted single strand, or the number of grafted strands. DNA-decorated colloids potentially exhibit extremely rich behaviour, since, in addition to modifying base sequencing, the colloidal

properties also can be altered. This tremendous number of possible choices makes DNA-linked assemblies one of the most versatile and promising new soft-matter materials and thus calls for theoretical [8–10] and numerical studies of these systems.

One of the key aims of such studies is the prediction of the three-dimensional self-assembled structure of these networked materials. It has been considered disappointing that most DNA-decorated colloidal dispersions form highly disordered aggregates, instead of crystal-like structures [4, 5, 11]. These materials are prone to form, in a reversible way, gels, i.e. disordered arrested states at low densities. In contrast to chemical gels, whose understanding has progressed much further due to the conceptual simplicity introduced by the infinite bond lifetime and the theoretical framework of percolation theory [12, 13], the thermo-reversible formation of colloidal gels in the absence of phase separation and crystallization is still an open problem in soft condensed matter [14, 15]. Recent studies have suggested that the generation of a thermo-reversible physical gel at low temperature T—i.e. conditions such that the bond lifetime is longer than the observation time and stresses can effectively propagate through the sample—is facilitated by limiting the preferred number of bonding neighbours (valency) [16, 17]. Such a constraint decreases the surface tension of a cluster aggregate, thereby destabilizing the phase separation process. In systems with spherically symmetric potentials, the bonding valency is only limited by packing considerations, and thus phase separation complicates the possibility of generating gel states. Molecular systems with highly directional interactions, such as occur in network forming liquids [18], are good candidates for atomic gels. Colloidal systems with patchy interactions and protein systems fall into the class of system for which the gel state is favoured. Novel biomaterials in which a specific (and small) number of complementary DNA strands are attached to a common centre [6, 7, 19] should naturally lead to a limited valency, and hence are an interesting group of potential gel-forming systems. The possibility of controlling both the number of arms (the valency) and the bonding energy (via the number of complementary sequences) makes these materials among the best candidates for checking the propensity to form gels and test recent explanations of physical gel formation [17], as well as for exploring the types and properties of self-assembled biomaterials.

Here we introduce a model designed to mimic the tetrameric DNA complexes recently synthesized and discussed in [19], but whose bulk behaviour has not yet been explored. The model is not designed to be chemically accurate, but should qualitatively reproduce the physical phenomena of DNA assembly. Using the model, we study the bulk behaviour of a system of many tetramers and find that, in an extremely narrow T range, the system forms a low-density disordered network via a thermo-reversible gel transition. In contrast with previously studied cases of thermo-reversible gelation [17, 21, 20, 15], the gel and percolation transitions of our model can be made arbitrarily close by exploiting the entropic contribution to bond formation made possible by the complementary DNA binding.

Each molecular unit of the model is composed of a tetrahedral hub tethering four identical DNA-like strands composed of eight connected monomers; we refer to this molecular unit as a tetramer. The ordering of the bases beginning from the tetrahedral core is A–C–G–T–A–C–G–T; A, C, G, and T are the standard abbreviations for the bases of DNA. Bases of type A bond only with type T, and type C bonds only with type G. We choose this sequencing since it offers the possibility of forming bonds between different tetramers in which all eight sites along a strand are paired. All pairs of sites interact via a purely repulsive potential obtained by truncating and shifting the Lennard-Jones (LJ) potential

$$V_{\mathrm{sf}}(r) = V_{\mathrm{LJ}}(r) - V_{\mathrm{LJ}}(r_{\mathrm{c}}) - (r - r_{\mathrm{c}}) \left.\frac{\mathrm{d}V_{\mathrm{LJ}}(r)}{\mathrm{d}r}\right|_{r=r_{\mathrm{c}}}. \tag{1}$$

The cutoff $r_{\mathrm{c}} = 2^{1/6}\sigma$, where σ is the LJ length parameter. Neighboring monomers (those

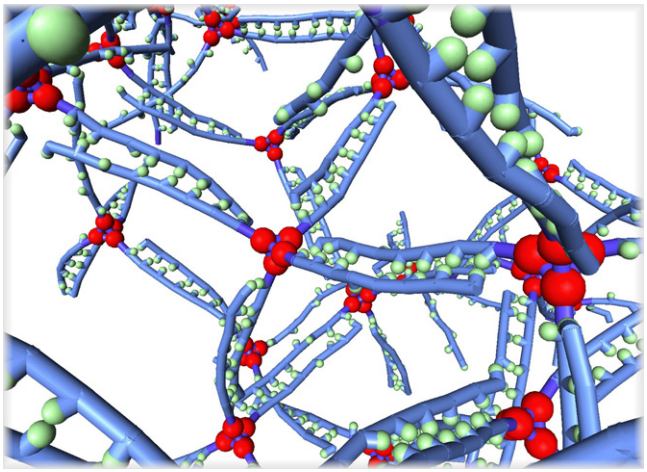

Figure 1. Snapshot of the simulation at low T where most base pairs are bonded. The large red spheres indicate the tetrahedral core of each DNA tetramer. The blue tubes indicate the bonds along a single strand of the DNA, and the small green spheres represent attractive base-pair force sites. The regular pairing of these green spheres shows the proclivity for complementary base pairs to attract each other.

along a strand and in the central tetrahedral hub) are connected via a finitely-extensible, non-linear elastic (FENE) anharmonic spring potential $V_{\rm FENE} = -k(R_0^2/2)\ln(1-(r/R_0)^2)$, where the bond strength $k = 30$, and the maximum bond extension $R_0 = 1.5$, as used in [22, 23] to study coarse-grained polymers. To model the characteristic rigidity of the DNA strands, we add a three-body potential of the form $k_\ell(1-\cos\theta)$, where θ is the angle defined by three consecutive monomers. A value of $k_\ell = 5$ allows for moderate flexibility of the strands, but prevents strands in the same tetramer complex from becoming entangled.

To simulate bond formation between complementary bases, each monomer along the strand has an additional 'bonding' force site that carries the information about the base type. Attractive interactions are included only between the bonding sites of complementary bases. The bonding sites are connected to the monomer core along the strand using the same FENE potential that links the strands together. The interactions between complementary bonding sites are modelled via an LJ potential as in equation (1), but the truncation distance $r_{\rm c} = 2.5\sigma_{aa}$ to include attractions. We choose $\sigma_{aa} = 0.35\sigma$ so that the bonding site is almost completely contained in the repulsive shell of the monomer core. This choice prevents the bonding site from connecting to more than one complementary base. Interactions between non-complementary bases are also given by equation (1) with $r_{\rm c} = 2^{1/6}\sigma_{aa}$, so that interactions are purely repulsive. Figure 1 shows a snapshot of the system when many base pairs are bonded. To keep the model parameters to a minimum, we do not include torsional terms which would be able to generate a helix geometry in the bounded double strands. We present our results in reduced units where length is in units of σ, time in units of $\sigma\sqrt{m/\epsilon}$, T in units $\epsilon/k_{\rm B}$ ($k_{\rm B}$ is Boltzmann's constant), and entropy is in units of $k_{\rm B}$.

We simulate the model via molecular dynamics calculations to generate data for the configurations and velocities of the constituent particles as a function of time. We use this information to examine the thermodynamic, structural, and dynamic properties; the dynamic properties are essential for detecting dynamic arrest phenomena. We study a system of $N = 200$ tetramers (a total of 13 600 force sites) in a box of length $L = 52.41$, resulting in a molecular number density $n = 1.39 \times 10^{-3}$. For this density, the approximate average

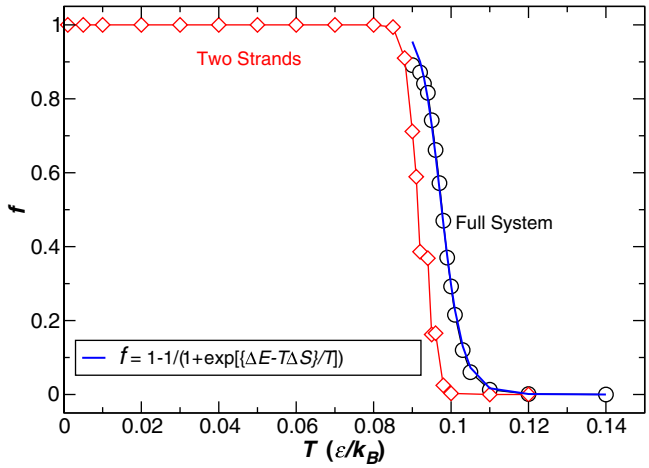

Figure 2. The fraction f of bonded DNA strands as a function of T. The figure shows f for the full system of 200 tetramers as well as for two isolated strands. The crossover from the unbonded to nearly fully bonded state is extremely sharp, but not discontinuous, and it is well described by the two-state model. We confirm the absence of hysteresis on reheating. We also performed a preliminary study on 64 tetramers which showed statistically identical behaviour, suggesting there are no significant finite size effects. In order for the plots to be comparable, both systems have a strand density of 0.0056.

distance between the centres of tetramers $\ell = n^{-1/3} = 8.96$. This separation is ideal for the formation of networks, since the distance between two bonded cores is $\approx$9, including the eight base pairs in the strand and the tethering monomers at the core. Each simulation is performed at a fixed density and T, and we control T using the Nose–Hoover method [24]. At each T we simulate two independent systems to improve our statistics. At the lowest T studied, our simulations extend to more than 10^7 time steps in order to access equilibrium behaviour. However, we point out that this lowest T may still exhibit modest ageing effects, but these will not affect our overall conclusions. To accelerate the overall speed of the simulations, we use a three-cycle velocity Verlet version of the rRESPA multiple time step algorithm, with the forces separated into rapidly varying bonded forces and more slowly varying non-bonded forces [24]. The time step for the bonded forces is 0.002.

To explore gel formation in this system, we study several different T values. Figure 2 shows the fraction f of bonded strands as a function of T. Since the attractive well between complementary base pairs is quite narrow, we say that a base pair is bonded if the energy between the pair is negative; we consider two strands to be bonded if at least half of the complementary base pairs of two strands are bonded. Figure 2 demonstrates that the range of the transition to a highly bonded state occurs over a narrow range $0.09 \lesssim T \lesssim 0.11$—only 2% of the energy of a single bond. A simple two-state relation $f = 1 - 1/(1 + e^{(\Delta E - T\Delta S/T)})$, in which the open and bonded states of the strand are attributed to an entropy change $\Delta S = 36$ and a energy change $\Delta E = 3.5$, and accurately describes the T-dependence of f. The large ΔS value shows that, in this model, bonding of strands freezes between four and five entropy units per base. This entropic contribution is responsible for the nearly first-order nature of the crossover to the bonded state, a feature not observed in recently studied physical gel models [17, 20, 15], the consequences of which will become apparent shortly. For comparison, figure 2 also shows $f(T)$ for two isolated strands. Despite the fact that the length of the strand is only eight monomers, a rather sharp sigmoidal shape characterizes the bond formation process.

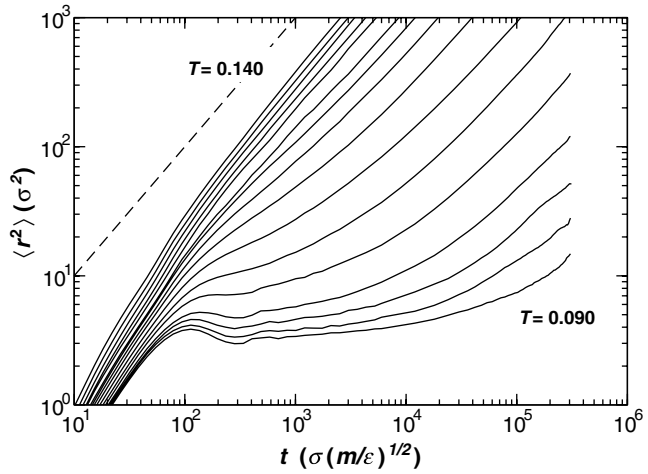

Figure 3. Mean-squared displacement $\langle r^2 \rangle$ of the tetramer centre of mass. The dotted line indicates the asymptotic linear behaviour of $\langle r^2 \rangle$ that is expected when tetramers are diffusive. The first four decades in time are not shown, since the dynamics at this time scale are trivial. From top to bottom, the T values we study are $T = 0.140, 0.120, 0.110, 0.105, 0.103, 0.101, 0.100, 0.099, 0.098, 0.097, 0.096, 0.095, 0.094, 0.093, 0.092, 0.090$.

While in the bulk system the competition between bonding arms prevents reaching the fully bonded state, the fully bonded ground state is easily reached at low T for the case of two isolated strands.

We next quantify the 'freezing' of the dynamic properties expected to occur once the gel state has formed. The mean-squared displacement $\langle r^2 \rangle$ of the tetramer centre of mass (figure 3) demonstrates that at high T, where $f \approx 0$, tetramers are able to diffuse with little hindrance. Over this narrow T range, $\langle r^2 \rangle$ becomes strongly hindered. Figure 4(a) shows the diffusion constant D calculated from the asymptotic behaviour of $\langle r^2 \rangle = 6Dt$ as a function of $1/T$. When few bonds are present, the T dependence of D is very weak. The slowing of the dynamics is intimately connected to the formation of bonds in the system; in the same narrow T region where bonding becomes significant, D dramatically decreases. Similar pronounced dependence of D on concentration for DNA dendrimers has been recently reported using particle-tracking experiments [25]. The nearly linear behaviour D at low T in figure 4(a) indicates a limiting Arrhenius behaviour, i.e. $\ln D \sim 1/T$. More importantly, we find a linear relation between D and $(1 - f(T))^4$ (figure 4(b)). Since f is the fraction of bonded strands, f also equals the probability that a strand is bonded. Thus $1 - f$ is the probability that a strand is unbonded, and so $(1 - f)^4$ is the probability that all four strands of a tetramer are unbonded. The quality of the comparison between D and $(1 - f)^4$ demonstrates that the variation of D at fixed density is controlled entirely by the fraction of fully unbound tetramers. This finding mirrors the behaviour recently found in a limited-valence model for colloidal gels [21]. Since $f(T)$ can be effectively described by the two-state model shown in figure 2, at low T

$$D = D_0[1 + \exp((\Delta E - T\Delta S)/T)]^{-4} \approx D_0 \mathrm{e}^{4\Delta S} \mathrm{e}^{-4\Delta E/T}, \tag{2}$$

which explains the observed low-T Arrhenius behaviour.

The low-T Arrhenius dependence of D means that motion will technically only cease in the limit $T \to 0$, but for practical purposes, the dynamics in the T window studied have already become much slower than any reasonable computational time. We call the T at which D can no longer be estimated the gel temperature $T_{\mathrm{gel}} = 0.092$, with $f(T_{\mathrm{gel}}) \approx 0.87$ and

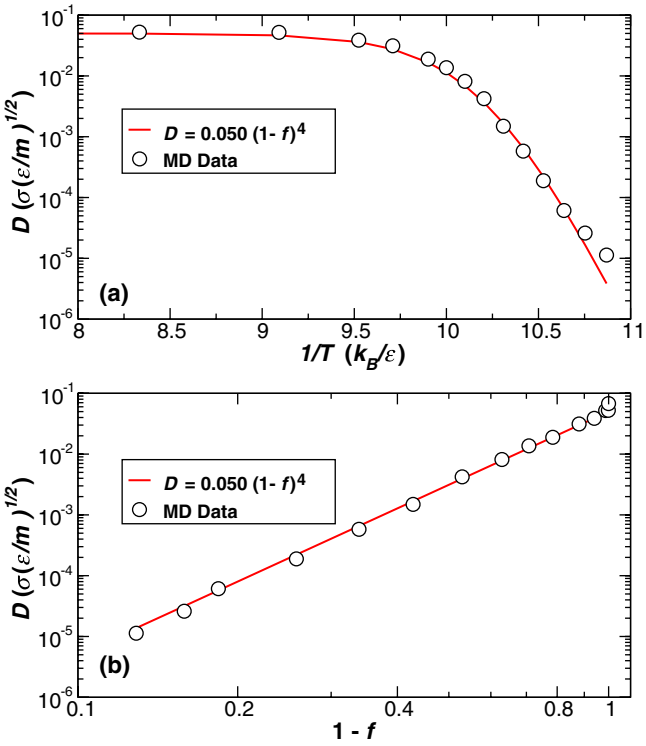

Figure 4. The relation of diffusion constant D to T and bonding fraction. (a) Arrhenius plot showing that D follows an Arrhenius law at low T. The line uses the two-state fitting form from figure 2. (b) Dependence of D on f; the line is the best fit between D and the raw data for $1 - f$, without using the two-state model. We find a power law relation $D = D_0(1 - f)^4$ over all T, where $D_0 = 0.050$. Hence D is controlled by the fraction of unbonded tetramers.

$D(T_{\rm gel}) \approx 1.1 \times 10^{-5}$. A much more restrictive definition of $T_{\rm gel}$, based on a diffusion coefficient ten orders of magnitude smaller—i.e. $D(T_{\rm gel}) \approx 10^{-15}$—would only move $T_{\rm gel} \approx 0.080$ (by extrapolation using equation (2)), a decrease of just 13%.

As we increase the length of strands, both ΔE and ΔS will increase, since they are proportional to the number of bonds between the strands. Combining this with equation (2) shows that increasing ΔS will cause D to decrease extremely rapidly over an even more narrow range of T, leading to the formation of an arrested state in a nearly discontinuous manner. Thus the sharpness of the crossover to the gel state can be tuned simply by changing the number of bases in the strand.

In recent model systems of physical gels [15, 20, 17] geometric percolation of clusters does not correlate with dynamic arrest, since at percolation, clusters break and reform continuously. Thus the T dependence of D is not strongly influenced by the crossing of the percolation locus. This is in sharp contrast with chemical gels (in which bonds form irreversibly), where gelation and percolation coincide. In our system, the close correlation between f and dynamics suggests that for sufficiently long strands the percolation of the DNA network is connected with the system's dynamic arrest, even though bonding is reversible. We locate the percolation transition using standard algorithms to partition tetramers into clusters of size s, identify spanning clusters, and evaluate the distribution of cluster sizes $n(s)$. Approaching percolation, $n_s \sim s^{-\tau}$, with $\tau \approx -2.2$, the theoretical value expected for random bond percolation [13]. We

find the percolation temperature $T_p \approx 0.099$. At this T, molecules are still able to diffuse but T_p is very near to $T_{gel} = 0.092$. For a chemical gel, $T_p = T_{gel}$ so that percolation ideas can be used to understand gel formation. Given the close correspondence between the length of stands and the activation energy for D, DNA-linked colloidal particles with controlled functionality make it possible to tune the distance between T_{gel} and T_p. The possibility of controlling in a reversible way the gel transition and the on–off character of the transition makes the DNA gels optimal biomimetic materials for delivery and release of host components.

In summary, we have shown the tendency for specifically sequenced DNA dendrimers to assemble into amorphous gel structures. In doing so, we demonstrated the close connection between the fraction of bonded strands and the dynamics of this new class of material. As a final comment, we recall that in the nanotechnology bottom-up approach, individual components are designed to assume particular tertiary structures with the aim of self-assembly into quaternary structures. Thus understanding the gel propensity of the designed nanostructure is fundamental for generating 3D structures with desired properties. Lastly, we stress the importance of understanding the conditions under which a system such as we have studied prefers to form a periodic crystal structure. Work is in progress in this direction.

We thank the NSF for support under grant number DMR-0427239; FS thanks also Miur FIRB and MRTN-CT-2003-504712. We thank J F Douglas and L W McLaughlin for valuable discussions, P Filetici for bringing reference [19] to our attention and the B U Center for Polymer Studies for hosting us during the initial stages of this research.

References

[1] Zhang Z L and Glotzer S C 2004 *Nano Lett.* **4** 1407–13
[2] Mirkin C A, Letsinger R L, Mucic R C and Storhoff J J 1996 *Nature* **382** 607–9
[3] Milam V T, Hiddessen A L, Crocker J C, Graves D J and Hammer D A 2003 *Langmuir* **19** 10317–23
[4] Kiang C-H 2003 *Physica* A **321A** 164–9
[5] Harris N C and Kiang C-H 2005 *Phys. Rev. Lett.* **95** 046101
[6] Paukstelis P J, Nowakowski H, Birktoft J J and Seeman N C 2004 *Chem. Biol.* **11** 1119–26
[7] Seeman N C 2003 *Nature* **421** 427–31
[8] Lukatsky D B and Frenkel D 2005 *J. Chem. Phys.* **122** 214904
[9] Lukatsky D B and Frenkel D 2004 *Phys. Rev. Lett.* **92** 068302
[10] Trachenko A V 2002 *Phys. Rev. Lett.* **89** 148303
[11] Biancaniello P L, Kim A J and Crocker J C 2005 *Phys. Rev. Lett.* **94** 058302
[12] Stauffer D and Aharony A 1994 *Introduction to Percolation Theory* (London: Taylor and Francis)
[13] Torquato S 2001 *Random Heterogeneous Materials: Microstructure and Macroscopic Properties* (New York: Springer)
[14] Rubinstein M and Dobrynin A V 1999 *Curr. Opin. Colloid Interface Sci.* **4** 83–7
[15] Kumar S K and Douglas J F 2001 *Phys. Rev. Lett.* **87** 188301
[16] Tanaka F and Stockmayer W H 1994 *Macromolecules* **27** 3943–54
[17] Zaccarelli E, Buldyrev S V, La Nave E, Moreno A J, Saika-Voivod I, Sciortino F and Tartaglia P 2005 *Phys. Rev. Lett.* **94** 218301
[18] Debendetti P G 1997 *Metastable Liquids* (Princeton, NJ: Princeton University Press)
[19] Stewart K M and McLaughlin L W 2004 *J. Am. Chem. Soc.* **126** 2050–7
[20] Dudowicz J, Freed K F and Douglas J F 1999 *J. Chem. Phys.* **111** 7116–30
[21] De Michele C *et al* 2005 *Preprint* cond-mat/0510787
[22] Grest G S and Kremer K 1986 *Phys. Rev.* A **33** 3628–31
[23] Bennemann C, Paul W, Binder K and Dünweg B 1998 *Phys. Rev.* E **57** 843–51
[24] Frenkel D and Smit B 2002 *Understanding Molecular Simulation* (San Diego, CA: Academic)
[25] Freedman K O, Lee J, Li Y, Luo D, Skobeleva V B and Ke P C 2005 *J. Phys. Chem.* B **109** 9839–42