Self-assembling DNA dendrimers: A numerical study

L. Largo
Francis Starr
Wesleyan University, fstarr@wesleyan.edu

F. Sciortino

Follow this and additional works at: http://wesscholar.wesleyan.edu/div3facpubs

Part of the Physics Commons

Recommended Citation

This Article is brought to you for free and open access by the Natural Sciences and Mathematics at WesScholar. It has been accepted for inclusion in Division III Faculty Publications by an authorized administrator of WesScholar. For more information, please contact dschnaidt@wesleyan.edu, ljjohnson@wesleyan.edu.
DNA is increasingly used as a specific linker to template nanostructured materials. We present a molecular dynamics simulation study of a simple DNA–dendrimer model designed to capture the basic characteristics of the biological interactions, where selectivity and strong cooperativity play an important role. Exploring a large set of densities and temperatures, we follow the progressive formation of a percolating large-scale network whose connectivity can be described by random percolation theory. We identify the relative regions of network formation and kinetic arrest versus phase separation and show that the location of the two-phase region can be interpreted in the same framework as reduced valency models. This correspondence provides guidelines for designing stable, equilibrium self-assembled low-density networks. Finally, we demonstrate a relation between bonding and dynamics, by showing that the temperature dependence of the diffusion constant is controlled by the number of fully unbound dendrimers.

I. Introduction

The use of biological cross-linking molecules to control bond formation between otherwise noninteracting particles offers a novel route to designing interconnected and highly organized materials on the molecular scale.\(^1\)\(^-\)\(^5\) Complementary “lock-and-key” binding molecules, which assemble due to the biological specificity, provide a versatile way of controlling interparticle binding. Among the possible choices of biological molecules, the specificity offered by complementary DNA single strands binding. Among the possible choices of biological molecules, specificity, provide a versatile way of controlling interparticle “key” binding molecules, which assemble due to the biological paradigm is the prediction of the three-dimensional self-assembled structure, the existence of ordered phases, the region of mechanical instability of the material (against composition fluctuations), and the propensity to form kinetically arrested states (gels or glasses). Ideally, through the design of the interparticle interaction, it should be possible to favor (or disfavor) crystallization, phase separation, or dynamic arrest. Indeed, the ability to fully exploit the rapid developments taking place in materials science requires not only an understanding of the equilibrium phases and their modifications with the external fields, but also an understanding of the regions in the phase diagram where nonergodic arrested states can interfere with assembly.\(^2\)\(^-\)\(^3\)

An additional valuable feature of using DNA strands for interparticle bonds is the possibility of controlling the maximum number of bonded neighbors, or valency. Controlling the valency is not a trivial task, given the propensity of most attractive spherically symmetric potentials to form highly packed structures with up to twelve neighbors. Valency can be controlled by decorating the surface of the particle with attractive patches, but materials.\(^1\)\(^-\)\(^9\) One of the key issues in the self-assembly paradigm is the prediction of the three-dimensional self-assembled structure, the existence of ordered phases, the region of mechanical instability of the material (against composition fluctuations), and the propensity to form kinetically arrested states (gels or glasses). Ideally, through the design of the interparticle interaction, it should be possible to favor (or disfavor) crystallization, phase separation, or dynamic arrest. Indeed, the ability to fully exploit the rapid developments taking place in materials science requires not only an understanding of the equilibrium phases and their modifications with the external fields, but also an understanding of the regions in the phase diagram where nonergodic arrested states can interfere with assembly.\(^2\)\(^-\)\(^3\)

An additional valuable feature of using DNA strands for interparticle bonds is the possibility of controlling the maximum number of bonded neighbors, or valency. Controlling the valency is not a trivial task, given the propensity of most attractive spherically symmetric potentials to form highly packed structures with up to twelve neighbors. Valency can be controlled by decorating the surface of the particle with attractive patches, but...
the patchy attraction range must be \( \approx 10\% \) (or less) of the particle diameter. It has been suggested that, in the absence of long-range repulsion, valency is the key ingredient to control the size of the region (in the density–temperature plane) in which phase separation is observed.\(^{(24,25)}\) Decreasing the valency stabilizes the system against concentration fluctuations and favors the possibility of forming, even at relatively low densities, homogeneous states held together by the percolating network of interparticle bonds. Moreover, patchy interactions that limit valency are also useful for the generation of complex self-assembled structures.\(^{(26,27)}\)

For a deeper understanding of the self-assembly process in DNA-coated particles, we study the structure and the dynamics of a minimal model recently introduced to mimic particles composed of a central core (which could be a molecule or a colloidal particle) to which several single-stranded DNA “arms” have been attached.\(^{(28)}\) Specifically, we focus on the tetrameric DNA complexes (four-armed oligonucleotide Ni(II)-cyclam-centered complexes) recently synthesized and discussed in ref \(^{(29)}\), but whose bulk behavior has not yet been experimentally explored. We investigate a large window of densities \( \rho \) over the relevant range of temperatures \( T \) to determine the region in which the system forms a supramolecular network, the region where the network becomes nonergodic, and the region in which network formation is precluded by phase separation; a schematic representation of these regions is shown in Figure 1, and the detailed results for this model are given in Figure 2. The details of Figure 2 will be developed throughout the manuscript.

In this manuscript, we characterize the network structure (in both real and reciprocal space), examine the connectivity of the system and map it to well-known percolation results, and evaluate the dynamical properties to elucidate the relation between bond formation, percolation, and dynamic arrest. Our results show that both the thermodynamic and dynamic features of this model are analogous to the properties of other tetrahedrally coordinated systems—indeed of their being molecular or colloidal. This provides further evidence that materials as different as the tetrahedral molecular fluids (like water and silica)\(^{(29–33)}\) and four-coordinated DNA-coated particles can be interpreted in the same general framework. Ultimately, this helps provide a reference for predicting self-assembled properties of particles with selective bonding.

### II. Model and Simulation Methods

The model we studied was designed\(^{(29)}\) to mimic the general features of the four-armed DNA dendrimer complexes that have been synthesized experimentally.\(^{(6)}\) Each arm consists of a single DNA strand with an even number of bases; the sequencing of the bases is chosen such that two identical strands will bond to each other in head-to-tail order. As a result, these dendrimers have the potential to naturally assemble into higher-order structures. The head-to-tail bonding is achieved by choosing a sequence where the bases of the second half of a strand are complementary to the bases of the first half, but in reverse order. For example, for the 8-mer strands we study, the first four bases in the strand are A-C-G-T. The complementary sequence to this is T-G-C-A, and so the overall sequence for the 8-mer is A-C-G-T-A-C-G-T. Reading the sequence backward, it is plain to see that it is the complement of the forward sequence. With this choice, it is also possible to have partial bonding of the outer four base pairs of the arms, due to the repetition of the A-C-G-T sequence.

---


\(^{(33)}\) Pratt, L. R. (Guest Editor) Chem. Rev. 2002, 102 (Thematic Issue Water).
Ideally, we would use a model potential that includes all the chemical details of DNA and which would explicitly account for electrostatic effects in the formation of DNA pairs. However, simulating a large system and following the self-assembly process presents significant computational challenges. Hence, we choose to avoid an explicit description of the solvent and of the charges, and we choose a simple, coarse-grained model of a DNA strand that captures the salient features of assembly made possible by the specificity of base pair interactions. The model is thus not designed to describe in a realistic way the characteristics of the DNA. The model has been described briefly in ref 28, which we expand on. The essential ingredients of the model are that each strand consists of an even number of monomers and each monomer carries a small “sticky spot” of a specific base type that will bond only with the complementary type. The small size of the sticky spot relative to the overall monomer size prevents a base from bonding with more than one complementary base, necessary to reproduce the selectivity of DNA base pair bonding.

We first describe in detail the potential we use for monomers of the DNA strand. Each strand has eight monomers, though the model is easily generalized to any number of monomers. All monomers (in the same or different strands) interact via the shifted-force potential

\[ V_{\text{sf}}(r) = V_{\text{LJ}}(r) - V_{\text{LJ}}(r_c) - (r - r_c) \frac{dV_{\text{LJ}}(r)}{dr} \bigg|_{r=r_c} \]  

where \( V_{\text{LJ}}(r) \) is the common Lennard-Jones (LJ) potential, and \( r_c \) is the distance at which the potential is truncated. This form of the potential ensures that both the potential and the forces are continuous at \( r_c \). For the monomers, we choose \( r_c = 2.16 \sigma \), where \( \sigma \) is the LJ length parameter; this \( r_c \) value is precisely the location of the minimum of the Lennard-Jones potential, and so the resulting potential is purely repulsive, as in the Weeks–Chandler–Andersen model. Hence, each monomer can be considered a soft sphere of diameter \( \sigma \). In the remainder of the text, we utilize reduced units where length is given in units of \( \sigma \) and energy in units of \( \epsilon \), the LJ energy parameter. Other derived units will be defined as they arise.

Connectivity of neighboring monomers in a strand is maintained by a finitely extensible, nonlinear elastic (FENE) anharmonic spring potential

\[ V_{\text{FENE}} = -\frac{kR_0^2}{2} \ln[1 - (r/R_0)^2] \]  

where the bond length \( k = 30\epsilon/\sigma^2 \) and the maximum bond extension \( R_0 = 1.5 \), as used in refs 40 and 41 to study coarse-grained polymers. We model the characteristic rigidity of the DNA strands by a three-body “bending” potential

\[ V_{\text{lin}} = k_{\text{lin}} (1 - \cos \theta) \]  

where \( \theta \) is the angle defined by three consecutive monomers, and a value of \( k_{\text{lin}} = 5\epsilon \) allows for moderate flexibility of the strands. At this level of description, we have a semiflexible polymer with only excluded-volume interactions.

The sticky spots account for base-specific attractive interactions; with any noncomplementary force site, the interactions will be purely repulsive. Each monomer carries one sticky site with an identity of A, C, G, or T that is bonded to the center of the monomer using the FENE potential (eq 2), with the same \( k \) and \( R_0 \). Sticky sites interact with noncomplementary sites and with the large monomers of the chain just described using the LJ potential (eq 1) with a diameter \( \sigma_{\text{sticky}} = 0.35 \); we chose \( r_c = 2.5\sigma_{\text{sticky}} \) so that all noncomplementary interactions are repulsive. This combination of the LJ and FENE potentials results in the center of the sticky site inside the diameter of the monomer, and so the “edge” of the sticky site just protrudes from the monomer, as illustrated in Figure 3a. This choice is important to prevent the formation of more than one bond per sticky site. Including attractions between complementary sites simply requires that we change the truncation of the potential; for complementary pairs of sticky sites, we choose \( r_c = 2.5\sigma_{\text{sticky}} \). At this level of description, the model can account for the zippering transition by which double-stranded DNA is formed (Figure 3b), as previously demonstrated. However, we note that we have not included interactions to mimic the helix spiraling of double strands.

To create four-armed dendrimer molecules, which we refer to as tetramers, we bond one end of each strand to a tetrahedral hub composed of four spheres of diameter \( \sigma \), as illustrated in Figure 3c. Monomers of the hub interact via the same repulsive potential as the monomers (eq 1), and bonds within the hub and between the hub and the strand are given by the FENE potential (eq 2). To maintain the orientation of the arms relative to the hub, we again use the three-body bending potential (eq 3). Specifically, for each vertex of the tetrahedral hub, we use eq 3 to prefer a linear orientation from the attached DNA monomer, through the vertex, and to each of the other neighboring vertices. The superposition of these linear potentials results in the arm having an orientation that is roughly normal to the opposing face of the tetrahedral hub.

We simulate 200 molecules (68 sites per tetramer, a total of 13 600 force sites) at six densities \( \rho = 0.01, 0.02, 0.03, 0.04, 0.06, \) and 0.07 at various \( T \) in the range 0.09 < \( T < 0.11 \), and include results for a seventh density \( \rho = 0.05 \) that have been previously reported. The density is defined as the number of monomers per unit volume (each molecule has 36 monomers). Therefore, \( \rho = 0.01 \) corresponds to a molecular density of 0.000 28. \( T \) is in units of \( \epsilon/k_b \), where \( k_b \) is Boltzmann’s constant. The range of \( T \) and \( \rho \) allow us to examine the relevant regions of the phase diagram of this system.

In order to evaluate both the static and dynamic properties of the system, we carry out molecular dynamics simulations. We simulate individual state points at a fixed \( \rho \) and \( T \); we control \( T \) via the Nose–Hoover method. 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 nonbonded forces. The time step for the bonded forces is 0.002, where time is in unit of \( \sigma\sqrt{m\epsilon} \). The time for equilibration of the system varies strongly with \( T \) but is only
is characterized by a soft depletion region around the origin, arising from the excluded volume interactions. The absence of peaks in $g(r)$ indicates an unstructured fluid. As $T$ decreases, we observe a clear peak at a separation $d \approx 9.6$, which defines the nearest-neighbor distance. We will use such a value of $d$ to define an appropriate scaled density to compare the phase diagram of Figure 2 with other tetrahedral models.\(^{29-35}\) On cooling, the first-neighbor peak becomes pronounced and a second peak develops at $r \approx \sqrt{3}d$, indicating that the system is forming a tetrahedral network. However, the structure of the system remains amorphous (noncrystalline).

To confirm the lack of crystal structure and the presence of tetrahedral order, we turn to $S(q)$, shown in Figure 5. Again, at high $T$, $S(q)$ is rather unstructured. On cooling, a main peak appears at $q \approx 0.8$, whose $q$-location does not changes with $T$. For $\rho \geq 0.3$, the tetrahedral network structure at low $T$ observed for $g(r)$ manifests itself as a shoulder (“pre-peak”) at $q \approx 0.6$. For sufficiently small $T$, the amplitude of the pre-peak, characteristic of the tetrahedral structure, grows. The absence of any Bragg peaks in $S(q)$ confirms the amorphous nature of the network.

For $\rho \leq 0.2$, one observes a sharp increase in $S(q)$ as $q \to 0$, an indication of an increase of the compressibility and of the development of inhomogeneities with a length scale on the order of the system size. The behavior of $S(q)$ for small $q$ is characteristic of an approach to a region of thermodynamic instability, which preempts the possibility of developing an homogeneous network at small $\rho$. To estimate the approximate location of the phase coexistence, we identify pairs of adjacent-state points which bracket the boundary between stable and unstable states. As a criterion for distinguishing single-phase equilibrium state points from phase-separating ones, we evaluate the value of $S(q_{\text{min}})$ where $q_{\text{min}} = 2\pi/L$ is the smallest wavevector available in the simulation study. State points for which $S(q_{\text{min}}) \leq 5$ are considered stable; states with $S(q_{\text{min}}) \geq 10$ are considered unstable. We thus bracket the stability boundary for this model by plotting in Figure 2 the last stable and first unstable states, both along isothers and along isochors. For $\rho \approx 0.3$, a weak increase of $S(q)$ at small $q$ is observed, but $S(q_{\text{min}})$ never goes beyond 1. For $\rho \approx 0.3$, critical fluctuations are not observed at all investigated temperatures. As shown in Figure 2, the phase separation region is limited to a relatively small region of the possible densities.

B. Potential Energy and Bonding. Since the only attractive intermolecular interaction is the complementary base-pair
interaction, the potential energy between base pairs provides an indication of the number of links in the system. In Figure 6, we show the potential energy \( E \) per base pair as a function of \( T \) for the different isochores studied. The energy decreases rapidly as we decrease the temperature for all the densities studied. In a very small interval of temperatures, roughly 2% of the energy can be associated with the energy needed to break a bond between strands, this energy should be nearly independent of density. Hence, we fix \( \Delta E = 4.0 \) and only allow \( \Delta S \) to vary, as shown in the inset of Figure 7. As expected, \( \Delta S \) decreases slightly with increasing density, since the overall volume is decreasing. In the two-state model, \( \Delta S \) is responsible for the abrupt change of \( \rho_b \) with decreasing \( T \). This also implies that a system with longer strands, i.e., where the loss of entropy on bonding will be larger, will form bonds in an even smaller interval of \( T \). Thus, width of the crossover could be made arbitrarily small by increasing the DNA strand length. The significant change in \( \Delta S/\Delta T \) in the bonding process (on the order of 5 per base pair or 2.5 per base) originates primarily from the localization of the base site, which is free to orient on the full solid angle in the unbonded state and is instead confined in a small angular range in the bonded one. In the present model, this is the main contribution, since there is no significant change in conformation of the strand between the bonded and unbonded states, due to our modeling of the bending energy. In real DNA, base pairs have comparatively less freedom, and hence our model overemphasizes this contribution to the entropy.

The two-state model also allows us to define the crossover to the predominantly bonded state by its inflection point.\(^44,46\) In the two-state model, the location of the inflection point coincides with the point where \( \rho_b = 0.5 \). Using eq 5 with the best fit parameters, we directly evaluate the inflection temperature as a function of density and plot it in the phase diagram of Figure 2. We will comment later on the relationship of this locus with the dynamics. The line shown in the diagram indicates a fit of the form \( \rho_b = 1 - \frac{1}{1 + e^{(\Delta E - \Delta S)/T}} \) (5).

To relate the connectivity of the system to its structural (and eventually dynamical) properties, an appropriate definition of a bond between distinct tetramers is needed. We say that a bond between base pairs of different arms is present when the pair base–base potential energy is negative; this definition works well, since there is a very narrow range of the attractive well. We say that a bond between two strands is present when the strands have four or more base-pair bonds formed.

We first evaluate the probability \( p_b \) that a strand of the tetramer is bonded. This probability is identical to the overall fraction of intact bonds in the system (i.e., number of intact bonds normalized by the maximum number of possible bonds). Similar to the behavior of the energy, Figure 7 shows that an arbitrary strand of the tetramers goes from almost zero probability of being bonded to another strand to nearly always being bonded in a small range of \( T \). A similar sharp behavior has also been observed experimentally for the melting transition of DNA-coated gold nanoparticles.\(^33\) The sharpness of the transition calls for a significant contribution of the entropy in the bond free energy. This sharp behavior is found for all the studied densities and can be rationalized, in a simple approximation, using a two-state model, in which the difference in bond free energy from an unbonded to a bonded state requires a change in both energy (\( \Delta E \)) and entropy (\( \Delta S \)). This hypothesis directly leads to

\[
p_b = 1 - \frac{1}{1 + e^{(\Delta E - \Delta S)/T}}
\]
The dashed line represents a power law with exponent $-2.2$ that corresponds to random percolation. Note that upon approaching the percolation transition the range of validity of the power law extends. The values of $T$ for the different densities are (a) 0.0965, 0.0970, 0.0975, 0.100, 0.105, 0.110; (b) 0.096, 0.0975, 0.100, 0.105, 0.110; (c) 0.095, 0.096, 0.097, 0.098, 0.099, 0.100, 0.110, 0.120; (d) 0.095, 0.097, 0.0975, 0.99, 0.100, 0.110.

Figure 9. Cluster size distribution $n(s)$ for several $\rho$ and $T$ values. The dashed line represents a power law with exponent $-2.2$ that corresponds to random percolation. Note that upon approaching the percolation transition the range of validity of the power law extends. The values of $T$ for the different densities are (a) 0.0965, 0.0970, 0.0975, 0.100, 0.105, 0.110; (b) 0.096, 0.0975, 0.100, 0.105, 0.110; (c) 0.095, 0.096, 0.097, 0.098, 0.099, 0.100, 0.110, 0.120; (d) 0.095, 0.097, 0.0975, 0.99, 0.100, 0.110.

proposed model of equilibrium polymerization.\textsuperscript{44,47} We find $E' = 5.6$, close to the energy needed to break a bond according to the two-state approximation. The temperature range over which we are able to test the equilibrium polymerization model is very small, but the behavior of the present set of data is consistent with this prediction. Should such a correspondence be robust, it would suggest that association leading complex network structures can be understood in the same framework as that of the simpler living polymer formation.

C. Percolation. We next consider whether the connectivity of the system can be understood in the framework of percolation theory.\textsuperscript{48,49} In Figure 8, we show the probability for a single DNA—dendrimer to have a specific number $n$ of arms bonded, with $0 \leq n \leq 4$. As expected from the energy, we find that in a small range of $T$, the tetramer goes from an almost zero probability of having an arm bonded to a significant probability to have all the possible bonds. We fit $P(n)$ with the function

$$P(n) = \frac{4!}{n!(4-n)!} p_b^n (1 - p_b)^{4-n}$$

that corresponds to random bond percolation with coordination 4.\textsuperscript{48} The agreement with eq 6 is excellent, and so we may interpret bonding as a random percolation process.

As a further test of the correspondence with random percolation, we examine the distribution $n(s)$ of clusters of finite size $s$. To evaluate $n(s)$, we examine all clusters and remove any that percolate the system. We define a percolating cluster as a cluster that spans the simulation box and is connected via the periodic boundary conditions. The periodic boundary condition is checked by replicating the simulation box in all directions of space and checking that in the replicated system the cluster is bigger than in the original nonreplicated system. Figure 9 shows that close to percolation $n_s \sim s^{-1}$ with $\tau \approx 2.2$. This exponent is exactly that expected for the random percolation process.

To locate the percolation line, we evaluate the fraction of configurations that contain a percolating cluster. Additionally, we evaluate $P_n$, the fraction of molecules belonging to the percolating clusters, which we find to be nearly identical to the fraction of percolating configurations (Figure 10). Since we find that in a very small range of $T$ the probability of finding a percolating cluster and the fraction of tetramers in the percolating cluster both go from 0 to 1, we can define the percolation threshold $T_{\text{percolation}}(\rho)$ as the $T$ where the probability of having a percolating cluster in an equilibrium configuration is 50%. We show the percolation line in the phase diagram (Figure 2). We find that the percolation locus occurs for $p_b$ values that go from ~0.42 to 0.32 as we go down in density, i.e., percolation occurs well before the completion of the network structure.

If bonds were irreversible (infinitely long-lived), the percolation line in the phase diagram would provide an indicator of a crossover to arrest, since the system will have a finite low-frequency shear modulus if static clusters percolate. However, our definition of percolation does not include any information about the lifetime of the spanning cluster. Since bonds can break and reform continuously, the percolation line cannot be taken as an indication of dynamical arrest. This is an important difference between physical and chemical gels; in the latter, the percolation line closely corresponds to the arrest line (gel line). At the same time, this model allows us to link physical and chemical gels, since by increasing the length of the strands we can make the percolation

\textsuperscript{(48) Stauffer, D.; Aharony, A. Introduction to percolation theory; Taylor & Francis: London, 1994.}  
line arbitrarily close to the dynamical arrest, as we will examine in the following section on dynamics.

IV. Dynamics

A. Bond Lifetime. In this section, we examine the dynamical properties of the system and how they are related to the connectivity of the network. As we discussed at the end of the previous section, the lifetime of the bonds is of particular interest, since this will determine to what degree we expect the percolation line to coincide with the arrest or gelation of the system. There are two ways one might evaluate the lifetime of a bond: (i) the mean time that an intact bond survives (mean first breaking time) or (ii) the characteristic time which an initially intact bond is permanently broken. The first measure is less relevant to the question of percolation, since bonds could potentially flicker on and off in a very short time scale but yet still be likely intact at some long time in the future. The second measure is more relevant, since it relates to the probability that an initially intact bond is intact at some later time, regardless of whether there were intervening intervals in which the bond was broken. Hence, we focus on this second measure.

To evaluate this measure, we consider the fraction of bonds \( f(t) \) that are present at some arbitrary initial time; note that \( f(t = 0) = p_0 \). We then evaluate \( f(t) \), the fraction of that same initial set that are also bonded at some later time \( t \). Formally, we can define a bond function

\[
h_{ij}(t) = \begin{cases} 1 & \text{strands } i \text{ and } j \text{ are bonded} \\ 0 & \text{strands } i \text{ and } j \text{ are not bonded} \end{cases}
\]

and the associated correlation function

\[
f(t) = \langle h_{ij}(t) h_{ij}(0) \rangle
\]

where the average is taken over all strand pairs and time origins.

In Figure 11, we show \( f(t) \), which demonstrates that the increase in the instantaneous fraction of bonds \( f(0) \) (already seen in Figure 7) is accompanied by an increase in the overall lifetime of the bond. The characteristic bond lifetime can be defined in several ways—for example, as the time at which \( f(0) \) has been reduced by one-half. To evaluate a characteristic time, we fit \( f(t) \) with a stretched exponential

\[
f(t) \approx A e^{-(t\tau)^\beta}
\]

and define the characteristic bond lifetime

\[
\tau \equiv \int_0^\infty f(t) \, dt = \tau_0 \Gamma(1/\beta)/\beta
\]

where \( \Gamma(\beta) \) is the Euler \( \Gamma \) function. We find that \( \beta \) is a weak function of \( T \), and for all state points, \( \beta \approx 0.7 \pm 0.1 \). The bond lifetime is almost \( \rho \)-independent, compared to the influence of \( T \), as shown in Figure 12. In other words, \( \rho \) only affects the number of bonds, but not the change in the rate of creation/destruction of these bonds with \( T \). The finite values of the bond lifetime, even for temperatures well below the percolation threshold, emphasizes that the network is transient and that bonds break and reform continuously. In this respect, percolation does not coincide with structural arrest in the system.

B. Diffusion and Isodiffusivity Lines. To better appreciate the correlation between the structural properties and the eventual dynamic arrest of the system, we next evaluate the mobility of tetramers by evaluating the diffusion coefficient as a function of \( T \) and \( \rho \). This will also allow us to extract lines of constant diffusion, or “isodiffusivity” lines. The location of these lines in the phase diagram indicates the eventual location of arrest and demonstrates some correlation with the percolation line and inflection of the two-state model.

To determine the diffusion coefficient \( D \), we calculate the mean-squared displacement \( \langle r^2(t) \rangle \) of the centers of mass of the tetramers (Figure 13). At short times, \( \langle r^2(t) \rangle \) is proportional to \( t^2 \), the so-called ballistic regime where the proportionality constant is the thermal velocity \( \langle v^2 \rangle \propto k_B T \). At very long times, \( \langle r^2(t) \rangle \) is linear in time, the so-called diffusive regime. At low \( T \), there is caging between the two regimes that results in an intermediate time window in which \( \langle r^2(t) \rangle \) is almost constant. Here, particles are confined by the bonds in an energetic cage (like the energetic cages of liquid water\(^{51,52}\)), with a localization length that can be inferred by the value of \( \langle r^2(t) \rangle \) at the plateau. As observed in glass-forming systems, this localization length is small compared to the first-neighbor distance. The anomalous oscillation peak observed in Figure 13c at low \( T \) is due to finite size effects which introduce a low-frequency cutoff in the vibrational density of states. The effect of cutting off the density of states shows up

only when the time scale for diffusion has significantly increased
and the system behaves in a solidlike manner for short time
scales.\textsuperscript{51,53} The diffusion coefficient \( D \) is defined as \( D = \langle r^2 \rangle / 6t \)
in the limit \( t \to \infty \). Figure 14 shows the dependence of \( D \) for all the
densities simulated along both isochoric and isothermal paths,
demonstrating that \( T \) dominates the dependence of \( D \), similar to
the behavior of \( r \). At high \( T \), where the number of bonds is \( \approx 0 \),
the tetramers are able to diffuse with little hindrance. In this high
\( T \) limit, diffusion is slowed down progressively by increasing
density, but only by about 1 order of magnitude in the investigated
range. A small decrease in \( T \) in the region where the network
forms has a dramatic effect in the slowing down of the dynamics.
Network formation results in a decrease of \( D \) by several orders
of magnitude. Figure 14 also shows that, when the network forms,
\( D \) is compatible with an Arrhenius dependence, i.e., \( D \sim e^{-E_D/T} \),
where \( E_D \) is the activation energy for diffusion. The origin of this
dependence and the value of \( E_D \) will become clear when we
examine the direct relationship between \( D \) and the fraction of
bonds in the system.

By making “cuts” at fixed values of \( D \) in Figure 14, we obtain
various loci in the \( T - \rho \) plane of constant \( D \), or isodiffusivity
lines (Figure 2). The progression of the isodiffusivity lines to
smaller \( D \) values provides an indication of the position and shape
of the eventual arrest line in the phase diagram. Figure 2 shows
that there is no significant change in the isodiffusivity lines when
they intersect the percolation line. Thus, the percolation line is
not directly related with dynamical arrest for this system, and
similar behavior has been found for other systems with finite
bond lifetimes.\textsuperscript{54,55} Additionally, there is no change in behavior
crossing the inflection point of the two-state model. Indeed, given
the Arrhenius behavior of \( D \), in a strict sense the system will only
arrest at \( T = 0 \); in a more practical sense, \( D \) will become extremely
small compared with experimental time scales for \( T > 0 \), but
lower than the lowest \( T \) we study. In either case, it is clear from
Figure 2 that the shape of the isodiffusivity lines is highly
correlated with the formation of the network as defined by either
of the static loci. Hence, these loci do provide useful estimates
of where the dynamics become influenced by the formation of
the network.

The phase diagram shown in Figure 2 provides a coherent
view of the behavior of the model in a large region of \( T \) and \( \rho \).

As alluded to before, recent studies of simple models of network-
forming liquids suggest the possibility of a universal scenario
controlled by the valency of the particles, in the absence of
isotropic attractive interactions. In this respect, all particles with
valency of 4 should be characterized by similar thermodynamic
and dynamic properties. A comparison of Figure 2 with Figure
9 of ref 34 shows that this is indeed the case. In all models for
tetrahedral particles studied, the region of thermodynamic
instability is observed only at low densities, delimited by the
density at which an unstrained network of fully bonded particles
can form—the optimal density of the network. In scaled density
(top axis in Figure 2), phase separation is encountered (in all the
models) only below \( \rho_{\text{scaled}} \approx 0.65 \). Above this value, no phase
separation is encountered on cooling, and the system can access
equilibrium states up to densities where the formation of a glass
due to packing constraints prevents equilibrium. In this window
of intermediate densities, kinetic arrest driven by the Arrhenius
increase of the bond lifetime provides the ultimate fate for the
system. In contrast, systems with simple isotropic attractions
(where the average number of bonded neighbors can go up to
ten) have no such density window, and the fate of the fluid
state at low \( T \) is either phase separation or glass formation.\textsuperscript{56}

C. Quantitative Relation of Diffusion and Network Forma-
tion. The previous study of this model\textsuperscript{28} found that, at the
preferred density for network formation, \( D \) appears to be
controlled by the fraction of completely unbonded tetramers.
Quantitatively, it was found that

$$D = D_0 (1 - p_b)^4$$

(10)

provides an accurate description of the diffusion coefficient. Since $1 - p_b$ is the probability that a bond is broken, $(1 - p_b)^4$ is the probability that all four bonds are broken, which corresponds to a completely free tetramer under the assumption of random bonding. Additionally, works with simple models of tetrahedral coordinated particles have shown the validity of the same relationship.\(^{35}\)

A possible model for the relationship between diffusion and bonding is that $D$ can be formally expressed as a sum over the contributions of particles with different numbers of bonds, that is

$$D = \sum_{n=0}^{4} D_n(T) P(n)$$

(11)

where $P(n)$ is the probability that an arbitrary dendrimer has $n$ bonds and $D_n(T)$ the diffusion coefficient for an $n$-bonded dendrimer. The strong assumption that only free particles ($n = 0$) are free to diffuse is equivalent to the $D = D_0 (1 - p_b)^4$ isochoric law, with the additional assumption that $D_b$ is weakly dependent on $T$. Figure 15 shows the relation between $D$ and $1 - p_b$ for the present model at all densities. With the exception of $\rho = 0.01$, we see that the $T$ dependence of $D$ is very satisfactorily described by eq 10. Finally, we point out that, since $p_b$ can be described by a two-state model, by combining eqs 5 and 10, we have

$$D = D_0 \left[ 1 + \exp \left( \frac{\Delta E - T \Delta S}{T} \right) \right]^{-1} \approx D_0 e^{4\Delta S} e^{-4\Delta E T}$$

(12)

as $T$ becomes small, which explains the observed low-temperature Arrhenius behavior of $D$. In this respect, $E_D$ is four times the bonding energy $\Delta E$. We also note that the prefactor of the Arrhenius law incorporate a large contribution arising from $\Delta S$. Additionally, since we know $D_b$, $\Delta E$, and $\Delta S$ from numerical fits, we can use eq 12 to estimate $D$ at arbitrarily low $T$. Inverting the relation, we estimate the temperature at which $D = 10^{-20}$, a value that would be small by experimental standards. We find that $D$ reaches this value at $T \approx 0.077$, $\approx 15\%$ below the lowest $T$ studied and weakly dependent on density—providing an estimate of the practical bound for arrest of the system.

V. Conclusions

We have presented a detailed study of the structure and dynamics of an off-lattice model of DNA-functionalized particles to provide an insight into their self-assembly. The model retains the essential ingredients of the DNA double-strand bonding (most importantly, the base-pair selectivity intrinsic in the specific A-T/G-C pairing). To facilitate computational study, the model coarse-grains most chemical details of DNA molecules. The DNA strands are modeled as a polymer whose beads are decorated by labeled sites mimicking the four different bases, with attractive interactions only with complementary bases. While the model is not designed to be chemically accurate, it captures the basic physics of the self-assembly process in these materials.

We have focused on an arbitrary choice of an eight-base DNA sequence, specifically the (A-C-G-T-A-C-G-T) sequence, which has the property of being complementary to the same sequence read backward (T-G-C-A-T-G-C-A). This makes it possible for pairs of strands grafted on different particles to selectively bind, promoting the formation of a three-dimensional network on cooling. More specifically, we have chosen to model a fourarmed dendrimer complex, which has been synthesized experimentally, in which the “hub” of the macromolecule is small compared to the DNA strands. Despite this specific choice, our results should be general to other DNA-derivatized particles, including gold and PMMA particles.

We have analyzed the structure of the fluid. In a small temperature interval, base pairing becomes favorable and the system crosses from an high-temperature fluid state (characterized by nearly all monomers) to a fully connected network of four-coordinated particles. At low densities, the driving force for bonding generates a phase-separation process, and the system separates into two coexisting phases (a gaslike phase and a connected liquidlike phase), as detected by the progressive growth of the density fluctuations and confirmed by visual inspection of the configurations.

For intermediate densities, the system does not phase-separate at any of the studied $T$. The driving force for phase separation is suppressed by the possibility of forming an unstressed network of four-coordinated particles. The thermodynamic behavior of the present model is analogous to the behavior observed in previously studied simple models of four-coordinated patchy particles.\(^{34,35}\) confirming that the reduced valency opens a region of densities where the system remains stable against phase separation to very low $T$.\(^{24,25,53,57}\)

The structure of the system (analyzed via the radial distribution function or the static structure factor) shows the clear pattern of tetrahedral coordinated systems, including the presence of a pre-peak in $S(q)$ and a ratio between the location of the first- and second-neighbor peaks of $g(r)$ significantly different from two. The process of network formation is well-described by random percolation theory. Close to percolation, the cluster size distribution is well-described by a power law with exponent $2.2$. On further cooling, a spanning cluster appears, which progressively incorporates all molecules in the system. Despite the presence of a percolating cluster, dynamics do not arrest, since the bond lifetime is finite, and hence the percolating structure is transient. We find that diffusion is mostly controlled by the particles that completely detach from the network and whose concentration is proportional to $(1 - p_b)^4$, in agreement with random percolation theory. This relation between diffusion and free particles has also been observed in other models of limited valence.\(^{35}\)

---

suggesting that the mechanisms of dynamics in network-forming materials have some common aspects. It is important to stress that, since the bond probability follows an Arrhenius law at low $T$, the $T$ dependence of the diffusional times also becomes Arrhenius on cooling. While this suggests that dynamics formally arrest only at $T = 0$ K, the significant contribution of the bond-formation entropy, amplified by the fourth power, gives rise to a significant prefactor in the Arrhenius law. As a consequence, in a very small $T$ interval, dynamics decreases by an order of magnitude of ten or greater, giving rise to an effective arrest transition at a finite temperature, which (in the present system) is not far from the percolation line. In this respect, while the low-temperature dynamics remains Arrhenius, the relative distance (in $T$) between percolation and dynamic arrest is significantly reduced as compared to simple tetrahedral models in which the bond formation has no significant entropy contribution.

By comparing the phase diagram of the model with that of simple tetrahedral patchy particles, we note a clear analogy between them. The location of the phase-separation region in scaled density, the shape of the percolation line, the shape of the isodiffusivity lines, and the $T$ and density dependence of the potential energy are identical in both models. The only significant difference between these two models is the width of the $T$ range over which $p_b$ crosses from 0 to 1. The large entropic contribution due to the collective nature of the bonding in the DNA strands causes the transition from independent units to a fully connected network in a very small $T$ window. In systems composed of tetravalent particles, it is possible to closely approach the fully bonded ground state without the interference of phase separation in a wide region of density. It is conceivable that, in this region, the slowing down of the dynamics is so pronounced that the system becomes stuck in a kinetically trapped nonergodic state, preempting the possibility of generating ordered structures—the analog of crystal states in atomic systems. Finally, the present study and previous work on models for tetrahedral interacting simple particles call attention to the role of the maximum number of bonds per particle as the key ingredient in the classification of the collective behavior of interacting systems and their self-assembly.