Reversible self-assembly of patchy particles into monodisperse icosahedral clusters

Alex W. Wilber and Jonathan P. K. Doye[a)]
Physical and Theoretical Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QZ, United Kingdom

Ard A. Louis[a)]
Theoretical Physics, University of Oxford, Keble Road, Oxford OX1 3NP, United Kingdom

Eva G. Noya
Departamento de Química-Física, Facultad de Ciencias Químicas, Universidad Complutense de Madrid, E-28040 Madrid, Spain

Mark A. Miller
Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom

Pauline Wong
Department of Radiology, University of Cambridge, Addenbrookes Hospital, Cambridge CB2 2QQ, United Kingdom

We systematically study the design of simple patchy sphere models that reversibly self-assemble into monodisperse icosahedral clusters. We find that the optimal patch width is a compromise between structural specificity (the patches must be narrow enough to energetically select the desired clusters) and kinetic accessibility (they must be sufficiently wide to avoid kinetic traps). Similarly, for good yields the temperature must be low enough for the clusters to be thermodynamically stable, but the clusters must also have enough thermal energy to allow incorrectly formed bonds to be broken. Ordered clusters can form through a number of different dynamic pathways, including direct nucleation and indirect pathways involving large disordered intermediates. The latter pathway is related to a reentrant liquid-to-gas transition that occurs for intermediate patch widths upon lowering the temperature. We also find that the assembly process is robust to inaccurate patch placement up to a certain threshold and that it is possible to replace the five discrete patches with a single ring patch with no significant loss in yield. © *2007 American Institute of Physics.*
[DOI: 10.1063/1.2759922]

I. INTRODUCTION

The remarkable ability of biological matter to robustly self-assemble into well-defined composite objects excites the imagination, suggesting that these processes could perhaps be emulated through the judicious design of synthetic building blocks.[1] Viruses provide a particularly inspiring example. As first shown 40 years ago for the cowpea chlorotic mottle virus,[2] and later for a wide variety of other species,[3–5] empty spherical virus shells (capsids) can be made to reversibly self-assemble from individual protein subunits (capsomers) *in vitro*, simply by changing solution conditions such as the *p*H. Although this process resembles micellar self-assembly, there is an important difference: in contrast to the polydisperse distributions that characterize micelles, the complete virus capsids are monodisperse.

Of those viruses with a roughly spherical structure, the vast majority are found to have icosahedral symmetry. Many of the structures observed were first predicted on symmetry grounds by Caspar and Klug.[6] More recently, Zandi *et al.*[7]

and Glotzer *et al.*[8] found that the most stable sizes of simple model systems correlated with many of the structures found in virus capsids.

By contrast, the mechanisms by which capsid self-assembly takes place are incompletely understood, though there has recently been considerable theoretical[9–20] and experimental[21–27] progress. Some common conclusions emerge from this body of work. First, one of the major potential obstacles to successful assembly results from a class of kinetic traps, in which a large number of partial capsids are formed but are then unable to proceed to completion because of a lack of remaining free capsomers. This situation can be avoided if the free energy barrier for nucleation is large, such that intermediates are formed only slowly and free capsomers remain plentiful.[16–23] Second, it has been shown that only weak capsomer-capsomer interactions are required for capsids to be stable, because of the large number of interactions present in a complete capsid.[19–24] It is worth noting that weak interactions will tend to result in a large nucleation barrier, and also allow easier breakdown of incorrectly formed intermediates. These factors together imply that in general weak interactions will favor successful assem-

---

a)Authors to whom correspondence should be addressed.

bly, so long as they remain strong enough for the complete capsid to be stable and for nucleation to occur on a reasonable time scale.

Understanding how self-assembly with viruslike control and fidelity could be achieved with synthetically produced subunits is an important goal for nanofabrication. New experimental techniques to create self-assembling systems are being rapidly developed.[1] For example, one particularly impressive recent example is the reversible formation of tetrahedra from four appropriately designed DNA strands.[28] For colloidal and nanoparticle building blocks, one of the most crucial requirements to allow increased control over assembly is the synthesis of anisotropic "patchy" particles. There has recently been a great deal of work in this area, with notable successes for both colloids[29–33] and polymer-coated nanoparticles.[34–36]

Both these experimental advances and the importance of understanding capsid formation have stimulated a number of simulation studies of the self-assembly of monodisperse objects. Rapaport[14] produced the first simulation study of capsid assembly and used a model consisting of trapezoidal particles with attractive sites which assemble into a viruslike structure. However, Rapaport's model uses several unphysical rules to assist in correct assembly, such as irreversible bonding if subunits come together in the correct conformation.

Subsequently, there have been a number of simulations that have achieved reversible self-assembly.[15,18,37–39] For example, Hagan and Chandler were able to assemble large 60-particle virus models by using as their basic "capsomer" units spherical particles with directional anisotropic interactions that are chosen to be complementary in order to help guide the particles into the right local relative orientations.[15] This feature may naturally be realized in protein-protein interactions but may be more difficult to implement in synthetic systems. Van Workum and Douglas found that viruslike assemblies (not necessarily monodisperse) formed from particles made up of three dipolar Stockmayer particles connected together to form triangles.[37] Zhang and Glotzer created anisotropic model particles by rigidly connecting spheres with different attractive potentials, and showed that these can form a rich variety of small cluster shapes and extended structures.[38,39] Most recently, Nguyen *et al.* performed extensive studies using a geometric model in which appropriately shaped particles with specific interactions reversibly assembled into icosahedral capsids.[18] They produced a phase diagram for the model and also observed incorrectly formed "monster" particles very similar to those seen with real viruses.

In this paper we describe a set of minimal models, single spheres with anisotropic patchy interactions, that exhibit reversible self-assembly into monodisperse cluster phases. Choosing such simple systems facilitates the systematic exploration of parameter space to uncover the optimal design rules for self-assembly, and helps untangle the roles of thermodynamic and kinetic factors. This model may offer insight into how synthetic nanoparticles and colloids could be designed to self-assemble, and while the model is too simplistic to accurately describe the interactions between the capsid proteins of viruses, we hope that the principles identified will also be of relevance to biological examples of monodisperse self-assembly.

II. METHODS

A. Model

Our model consists of spherical particles with a number of patches whose geometry is specified by a set of patch vectors. The repulsion is based on the isotropic Lennard–Jones potential

$$V_{LJ}(r) = 4\epsilon\left[\left(\frac{\sigma_{LJ}}{r}\right)^{12} - \left(\frac{\sigma_{LJ}}{r}\right)^{6}\right], \qquad (1)$$

but the attraction is modulated by an orientationally dependent term, V_{ang}. Thus, the complete potential is

$$V_{ij}(\mathbf{r}_{ij}, \Omega_i, \Omega_j) = \begin{cases} V_{LJ}(r_{ij}) & r < \sigma_{LJ} \\ V_{LJ}(r_{ij})V_{ang}(\hat{\mathbf{r}}_{ij}, \ \Omega_i, \ \Omega_j) & r \geq \sigma_{LJ}, \end{cases} \qquad (2)$$

where Ω_i is the orientation of particle i. V_{ang} has the form

$$V_{ang}(\hat{\mathbf{r}}_{ij}, \Omega_i, \Omega_j) = G_{ij}(\hat{\mathbf{r}}_{ij}, \Omega_i)G_{ji}(\hat{\mathbf{r}}_{ji}, \Omega_j), \qquad (3)$$

$$G_{ij}(\hat{\mathbf{r}}_{ij}, \Omega_i) = \exp\left(-\frac{\theta_{k_{min}ij}^2}{2\sigma^2}\right), \qquad (4)$$

where σ is the standard deviation of the Gaussian, θ_{kij} is the angle between patch vector k on particle i and the interparticle vector $\mathbf{r}_{ij}$, and k_{min} is the patch that minimizes the magnitude of this angle. Hence, only the patches on each particle that are closest to the interparticle axis interact with each other.

When using this potential in our Monte Carlo simulations, we truncate and shift the potential using a cutoff distance of $3\sigma_{LJ}$ in order to improve the computational efficiency. To maintain the isotropic nature of the repulsion, we also adjust the distance at which we start to include the angular modulation so that it corresponds to where the cut-and-shifted Lennard–Jones potential passes through zero.

Previously, this potential has been used to study the crystallization of patchy particles.[40,41] Here, the target structure we choose to try to assemble is the 12-particle hollow icosahedron [Fig. 1(b)]. This structure is sufficiently complex to allow interesting behavior to be observed, but forms easily enough for large numbers of simulations of acceptable length to be carried out. It is also analogous to the structure of "$T = 1$" viruses.

In order to design the patchy particles so that they self-assemble into hollow icosahedra, the natural choice for the patch positions is to place them such that they point directly at the neighboring particles in the target structure. The resulting particle design, along with an assembly of twelve such particles into the target structure, is shown in Fig. 1(b). Having chosen the patch positions, the only parameter in the potential that remains to be optimized is then the patch width σ.
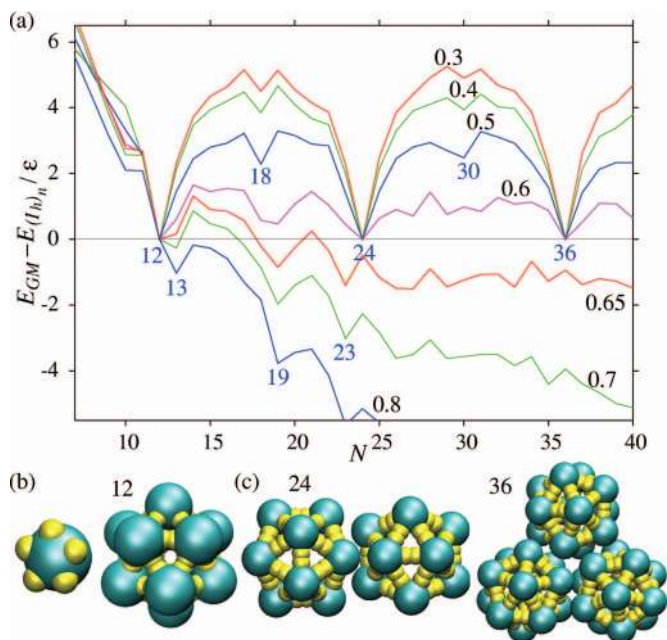
FIG. 1. (Color) (a) The size dependence of the energy, E_{GM}, of the global minimum for different values of σ. Each line is labeled by the value of σ (measured in radians), with particularly stable cluster sizes also labeled. The energy zero, $E_{(I_h)_n}$, is a fit to the energies of the assemblies of complete icosahedra possible at $N=12$, 24, 36, and 48 [some of which are shown in (b) and (c)] at the appropriate value of σ. Hence, negative values indicate that it will never be energetically favorable for the system to form monodisperse icosahedra. (b) A single particle designed to form a hollow icosahedron and a complete icosahedral cluster. The yellow spheres represent the positions of the patch vectors. (c) The minimum-energy structures for 24 and 36 particles for $\sigma \le 0.6$. The radius of the particles is reduced compared to (b) to allow the structures to be seen more clearly.

B. Monte Carlo simulation

We simulate the system in the canonical ensemble using a Metropolis Monte Carlo algorithm where the allowed move types are translations and rotations of individual particles. The translational moves are randomly chosen from a small cube centered on the selected particle. Rotational moves make use of a quaternion representation of the particle's orientation, which is modified by the addition of a smaller random quaternion and then renormalized.[42]

This algorithm does not directly simulate physical trajectories and is more often used to obtain thermodynamic averages of properties of a system. However, while the dynamics generated are not strictly rigorous, the use of small translational and rotational steps will mimic diffusive dynamics.[43–45] The advantage of this technique is that the derivatives of the potential are not needed, which allows the simulations to run very efficiently.

In the analysis of the results it is useful to assign the particles to clusters. The condition used is that two particles are considered bonded if their interaction energy is less than -0.4ϵ. Two particles are then considered to be in the same cluster if they are joined by a continuous chain of bonds.

As the above Monte Carlo simulations generally do not reach equilibrium, to determine some of the thermodynamic properties we also used the umbrella sampling technique.[46,47] In umbrella sampling, rather than generating configurations with the Boltzmann distribution, configurations are instead sampled according to the distribution $\exp\{-\beta[\mathcal{V}+w(Q)]\}$, where $w(Q)$ is a weighting function that is a function of an order parameter Q. Generally, the reason for introducing the weighting function is to artificially reduce the free energy barrier between different (meta)stable states of a system, hence facilitating interchange between these states and allowing the system to approach equilibrium. Canonical averages are simple to obtain from a simulation of such a non-Boltzmann (nB) ensemble using the expression

$$\langle B \rangle_{NVT} = \langle B \exp[\beta w(Q)] \rangle_{nB}, \tag{5}$$

where B is some generic property of the system.

In our case, we wish to locate the temperature at which the system becomes most stable as a collection of icosahedra, and so we need to overcome the free energy barrier for the formation of this structure. As the icosahedra only very weakly interact with each other, it is sufficient, to a first approximation, just to consider a 12-particle system. For an order parameter we simply use the number of particles in the largest cluster present. We wish to choose the weight function so that the system spends an approximately equal amount of time at each value of Q. To achieve this we use an iterative scheme in which $w(Q)$ is updated at the end of each of a series of simulations.

III. RESULTS

A. Energetics

A minimal requirement for self-assembly is that the system of monodisperse clusters is energetically most stable, i.e., a system of n icosahedra must be lower in energy than the lowest-energy single $12n$-particle cluster. We check for what range of σ this requirement holds, by performing global optimization to find the lowest-energy configurations of our system as a function of the number of particles. The global optimization algorithm we used was the basin hopping algorithm,[48] which has proved to be particularly successful for clusters.[49] Figure 1 shows that for wider patches the global minima correspond to single compact clusters with magic numbers as for the isotropic Lennard–Jones potential ($N = 13, 19, 23, \ldots$).[50] However, for sufficiently narrow patches ($\sigma \le 0.6$) the most stable sizes occur at $N = 12, 24, 36, \ldots$, and do indeed correspond to packings of $1, 2, 3, \ldots$ discrete icosahedra [Fig. 1(c)], as desired. (Note that as the icosahedra weakly interact, the discrete icosahedra are touching in the global minimum configuration.) The energy landscapes for such magic number clusters are likely to have a single-funnel topography.[51,52]

B. Self-assembly

To further investigate the self-assembly behavior of our system, we performed Monte Carlo (MC) simulations in the canonical ensemble. Figure 2 illustrates a typical trajectory, and shows that, given the right conditions, the patchy particles are able to reversibly assemble into a monodisperse set of hollow icosahedral clusters. At the beginning of the simulation, disordered clusters form rapidly [Fig. 2(a)]. These clusters then soon begin to order. For example, five-coordinate vertices become visible, and many of the particles
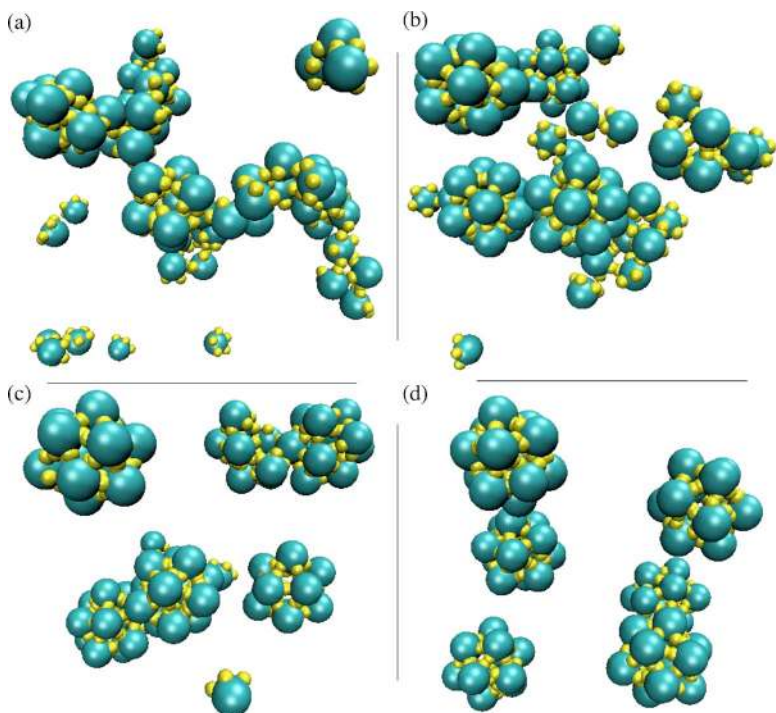
FIG. 2. (Color online) Snapshots of a system of 72 particles assembling into six complete icosahedra at σ =0.45 and $T=0.14\epsilon k_B^{-1}$ after (a) 3000, (b) 8000, (c) 60 000, and (d) 250 000 MC cycles.

already occupy their correct positions in partially formed icosahedra [Fig. 2(b)]. The slowest process is then the diffusion of the few remaining monomers to the empty sites in the clusters [Figs. 2(c) and 2(d)].

The simplicity of our model allows us to map out in detail the assembly behavior, thus enabling the optimal patch width and temperature to be identified. Figure 3 summarizes results from 480 simulations at different combinations of σ and T, where each simulation was started from a disordered configuration generated at high temperature. Although this plot is for a particular density and after a certain number of MC moves, the generic features are not sensitive to these parameters (for the dependence on density see Sec. III E).

The key feature of Fig. 3 is that there is a limited range in the space of patch width and temperature where self-assembly occurs efficiently. The maximum yield of 88% occurs at σ=0.45 and $T=0.14\epsilon k_B^{-1}$. The general patterns shown in Fig. 3 are relatively insensitive to simulation time, although in general yields increase with time. For example, we have observed yields as high as 98% after 800 000 MC cycles.

To understand the assembly kinetics summarized in Fig. 3, we first examine the effect of temperature at the optimal σ. There is a clear maximum in the number of icosahedra formed, with the number dropping to zero at high and low temperature. The upper limit is thermodynamic because at high temperatures the stable phase of the system is a gas of monomers. The lower limit is kinetic and arises because at low temperatures the system does not have enough thermal energy to escape from incorrect arrangements of the particles, and so large low-density kinetic aggregates grow. These two limits are clearly visible from the mean cluster size plotted in Fig. 3(b).

A snapshot of a kinetic aggregate is shown in Fig. 4(a). This structure is typical of those formed at higher σ and low temperature. The relatively wide patches allow significant rearrangement of the clusters, and it can be seen that there is a large degree of local order, with most particles forming strong interactions through most or all of their five patches. The diameters of the "strings" in the network are roughly equal to the diameter of an ideal icosahedron. However, the low temperature prevents complete rearrangement to form icosahedra. For lower σ and low T the frustration of the system is much more severe and most particles only manage to form two or three strong interactions [Fig. 3(c)]. The resulting network does not coalesce into dense clusters like those in Fig. 4(a), and the "threads" of the network are only one particle in diameter. At very low σ very few interactions are formed at all and only small clusters are found.

Having a temperature window where ordering can occur such as that observed here is common. For example, in the protein folding community it has been argued that increasing the ratio T_f/T_g (T_f is the "folding" temperature below which the native state becomes most stable and T_g is the "glass transition" temperature below which the dynamics becomes too slow for folding to occur) enhances the ability of a protein to fold.[53] Similarly, here we find that the optimal value of σ occurs roughly where there is the largest difference between T_{clust}, the temperature below which the clusters become thermodynamically most stable, and T_g, below which ordering is kinetically hindered.

The effect of reducing σ from its optimal value is to narrow the temperature window over which icosahedra form. First, T_{clust} decreases as the patches become narrower, because the vibrational entropy of the icosahedral clusters decreases. Second, the potential becomes "stickier" as the patches narrow, and so the glass transition temperature T_g below which the particles become trapped in large nonequilibrium aggregates initially increases. Consequently, for σ ≤ 0.1 icosahedra are virtually never seen.
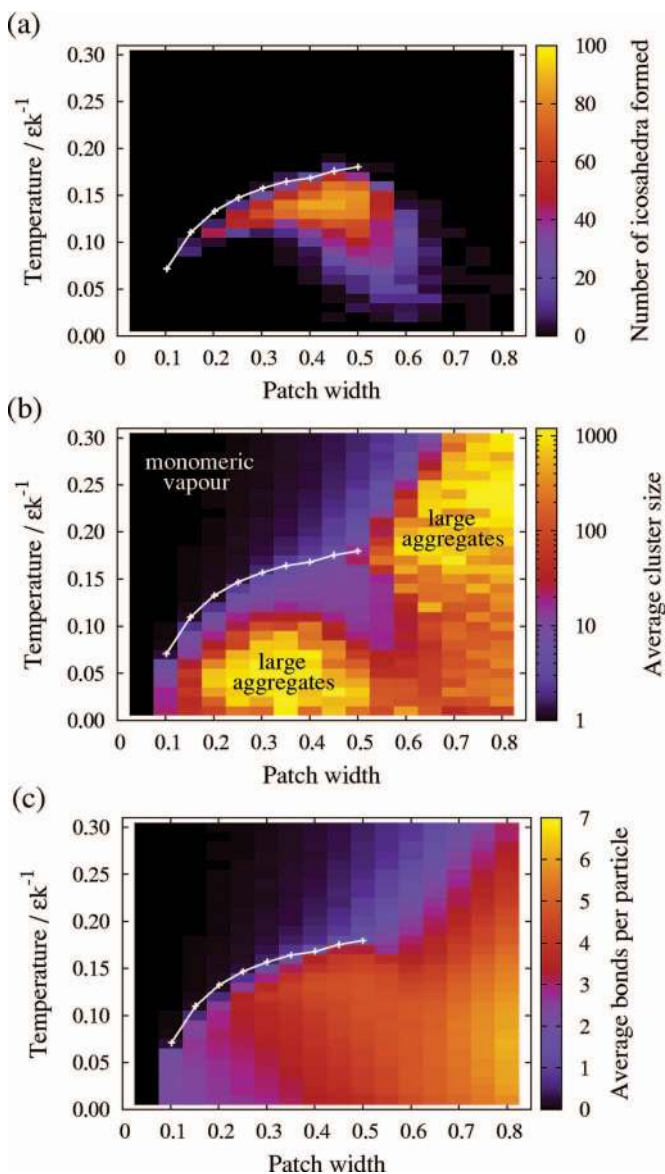
FIG. 3. (Color) (a) The number of icosahedra formed, (b) the mean cluster size (averaged over particles), and (c) the average number of bonds in which a particle participates after 80 000 MC cycles as a function of the patch width σ (measured in radians) and the temperature for 1200 particles at a number density of $0.15\sigma_{LJ}^{-3}$. The white lines show the thermodynamic transition temperature T_{clust} for the transition from icosahedra to a gas of monomers (see Sec. III C).
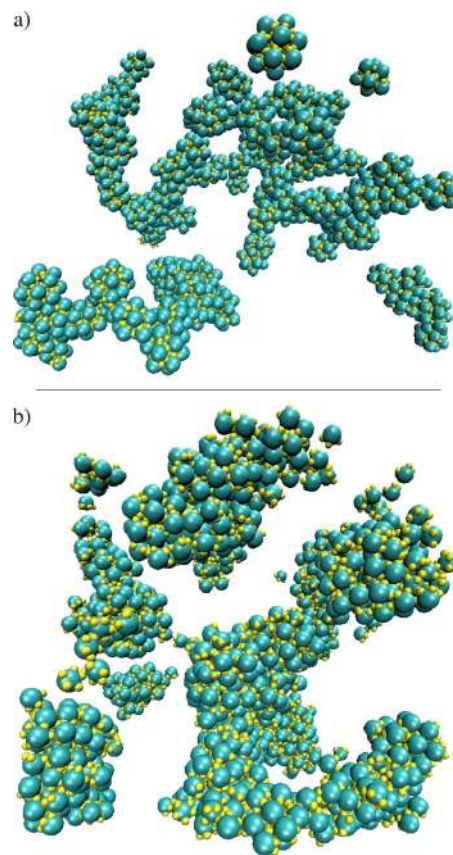

FIG. 4. (Color online) Typical aggregate structures formed at nonoptimal temperatures and patch widths. (a) Stringlike kinetic aggregates formed at $T=0.04\epsilon k^{-1}$ and $\sigma=0.6$. (b) Thermodynamic liquidlike aggregates formed at $T=0.18\epsilon k^{-1}$ and $\sigma=0.65$.

first condenses from a monomeric gas to a bulk liquidlike phase. However, at lower temperature it forms a gas phase again, but now made up of weakly interacting icosahedral clusters. This reentrant lower liquid-to-vapor transition is driven by the lower energy of the clusters, which overcomes the higher entropy of the liquid. Preliminary parallel tempering simulations have confirmed this unusual scenario. The probable implication is that there is a closed-loop liquid-vapor coexistence line in the phase diagram for this system. This will be explored further in future work.

C. Thermodynamics

In order to obtain estimates for T_{clust}, the temperature below which the system is most stable as a collection of icosahedra, we made use of the umbrella sampling scheme described in Sec. II B. The resulting heat capacity plots obtained for a selection of patch widths are shown in Fig. 5(a).

The peaks in the heat capacity signal the transition from a gas of monomers to a gas of icosahedra. As an aside, we should note that this transition is not a phase transition, because it is fundamentally a transition associated with a finite cluster, e.g., the heat capacity peaks will retain the same finite width, no matter how large the system is.

The positions of the peaks provide estimates for T_{clust}, which are shown superimposed on Fig. 3(a). It is interesting to note that relatively little supercooling is required for the icosahedra to begin to assemble. This is confirmed in Fig.

Figure 1 shows that beyond a certain σ, the icosahedra are no longer the most stable particle configuration. This is mirrored in Fig. 3(b) by the large disordered clusters that form for larger σ. At these values of σ and density the system phase separates into one or more liquid droplets coexisting with a vapor. A typical configuration from this region of the phase diagram is shown in Fig. 4(b). Note that the liquid droplets are not particularly spherical. The system has a low surface tension because most patches for the particles on the surface of the droplets are pointing inwards. As a result large shape fluctuations in the droplets are observed.

The interaction between the clustering transition and the formation of a liquid phase gives rise to a particularly intriguing form for the phase diagram of this system. For example, as the temperature is lowered for $\sigma \approx 0.55$, the system
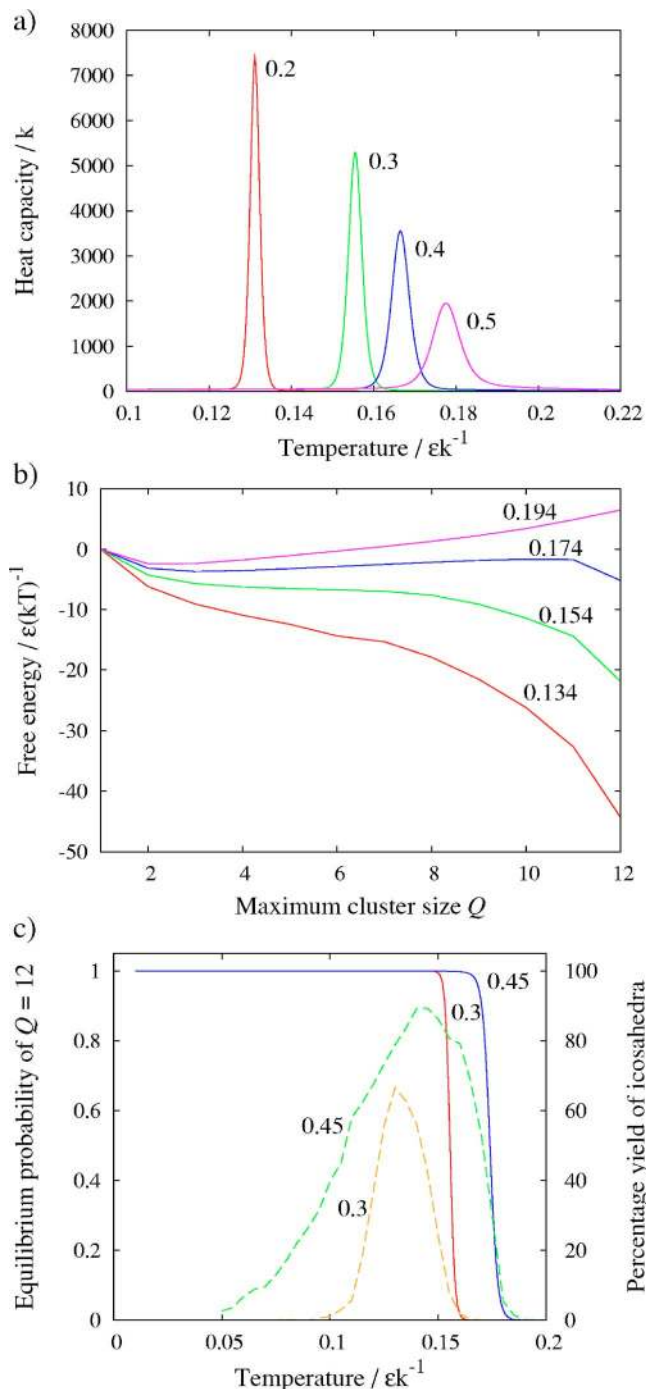
a)

b)

c)

FIG. 5. (Color online) Some thermodynamic properties of a 12-particle system. (a) The heat capacity as a function of temperature for different values of σ. Each line is labeled by the value of σ (given in radians). (b) The free energy associated with the order parameter Q plotted at four different temperatures, all at a patch width of $\sigma=0.45$. Each line is labeled by the temperature, with $T=0.174$ being the value of T_{clust} for this patch width. For clarity, we have chosen the energy zero so that $A_L(1)=0$. (c) Equilibrium probability of $Q=12$ as a function of temperature (solid lines). For comparison the yield of icosahedra obtained after 80 000 MC cycles in self-assembly simulations (dashed lines) is included. Each line is labeled with the value of σ (given in radians).

5(c) where we compare the equilibrium probability of a particle being in an icosahedron to that obtained in our MC simulations of the self-assembly. In particular, when the icosahedra first become stable as the temperature is de-
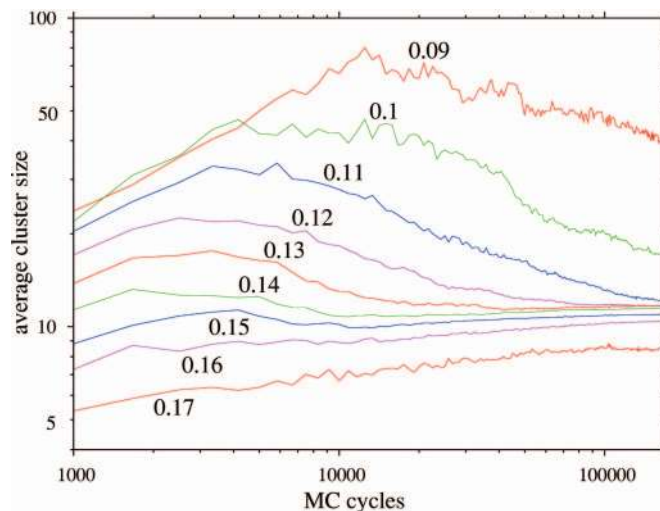

FIG. 6. (Color) The mean cluster size as a function of number of MC cycles at $\sigma=0.45$ and at different temperatures (as labeled) for a system of 1200 particles at a number density of $0.15\sigma_{\text{LJ}}^{-3}$. Each line is an average over ten simulations.

creased and their equilibrium probability is still low, the yield in our self-assembly simulations mirrors, to within statistical error, the equilibrium results.

The free energy profiles, $A_L(Q)$, can be obtained straightforwardly from the umbrella sampling simulations, through the relation

$$A_L(Q) = A(T) - kT \ln p(Q), \qquad (6)$$

where $A(T)$ is the Helmholz free energy. A series of profiles are plotted for the optimum patch width in Fig. 5(b). The change in the relative stability of the icosahedra compared to the vapor as the transition is crossed is apparent and for sufficiently low temperature the profile becomes barrierless. At T_{clust}, the free energy barrier is of the order of 2 kT. This relatively low value suggests that overcoming the nucleation barrier is not rate limiting and helps to explain why the yield of icosahedra initially follows the equilibrium line as the temperature decreases from above T_{clust} [Fig. 5(c)]. However, as the equilibrium population of icosahedra increases, the yield from our self-assembly simulations begins to deviate from the equilibrium line, initially because the time scale for diffusion of free particles to the correct sites on incomplete clusters begins to limit the yield.

D. Mechanisms of assembly

Further information about the mechanisms of self-assembly can be gleaned from Fig. 6 which shows how the average cluster size evolves with MC time for different temperatures at the optimal patch width. One striking feature is that for $T \lesssim 0.14\epsilon k^{-1}$ the average cluster size goes through a maximum before decreasing toward 12. The pathway to cluster formation is through larger disordered clusters, which then "bud off" icosahedra, rather than through the growth of icosahedra from smaller units. When an icosahedron with all the patches correctly oriented forms within a larger cluster, it will only interact very weakly with the rest of the cluster, and so will be able to escape relatively easily. The height of the

maximum in the average cluster size increases with decreasing temperature because the rate of cluster growth increases compared to the rate at which icosahedra are annealed out.

These results indicate therefore that the self-assembly of icosahedra is facilitated by the formation of a metastable intermediate phase, and the liquid phase, which we previously noted is stable at larger σ, plays an important role, even when it has disappeared from the equilibrium phase diagram. This system therefore provides an example of Ostwald's step rule[54] where nucleation initially leads to a metastable phase. For example, protein crystallization can be enhanced by the formation of a metastable protein-rich phase from which the crystal can more easily nucleate.[55,56]

Figure 7 illustrates some of the changes in assembly mechanism as a function of temperature and patch width. In general, conditions of high temperature and low σ favor a mechanism in which capsids are built up by a stepwise addition of monomers and small clusters. In Fig. 7(a) the cluster size distribution is dominated by monomers and small intermediates after 80 MC cycles. Over the course of the simulation the population of particles shifts steadily into icosahedra, with almost no clusters larger than 12 particles being formed. Conversely, conditions of low temperature and high σ (close to the range of parameters where there is a stable liquid phase[57]) favor the formation of large clusters, which then anneal to icosahedra. Figure 7(b) shows the rapid formation of very large clusters at the start of the simulation. The average cluster size then shrinks back down by the "budding" mechanism mentioned earlier until the majority of particles are in icosahedra. It is interesting to note that at long times the large clusters that are left preferentially adopt sizes that are magic numbers for the isotropic Lennard–Jones potential, e.g., the peak at $N=19$ in Fig. 7(b) that corresponds to two interpenetrating centered icosahedra.[50] At the optimal conditions for assembly [Fig. 7(c)], a combination of the two mechanisms seems to operate.

One effect of the availability of this budding mechanism is to increase the range of temperature over which icosahedra assemble. For example, it is noticeable from Fig. 5(c) that as σ decreases and the system moves further from the region of parameter space where the liquid phase is stable, the temperature range for which there are significant yields decreases significantly.

The budding mechanism of self-assembly is very different from that so far observed in experimental[21,26,58] and theoretical (both kinetic models[16,58] and direct simulations[15]) work on viruses, where direct nucleation of the virus capsid by stepwise addition of individual capsomers is the norm. We suspect that the greater specificity of the interactions in real viruses both makes the formation of large aggregates less likely under conditions where the complete capsids are most stable and prevents the easy rearrangement of subunit particles in any large aggregates that form.

E. Dependence on density

All of the results described thus far are for simulations at the same density, $0.15\sigma_{LJ}^{-3}$, which raises the question of how the assembly behavior depends on particle density. Figure
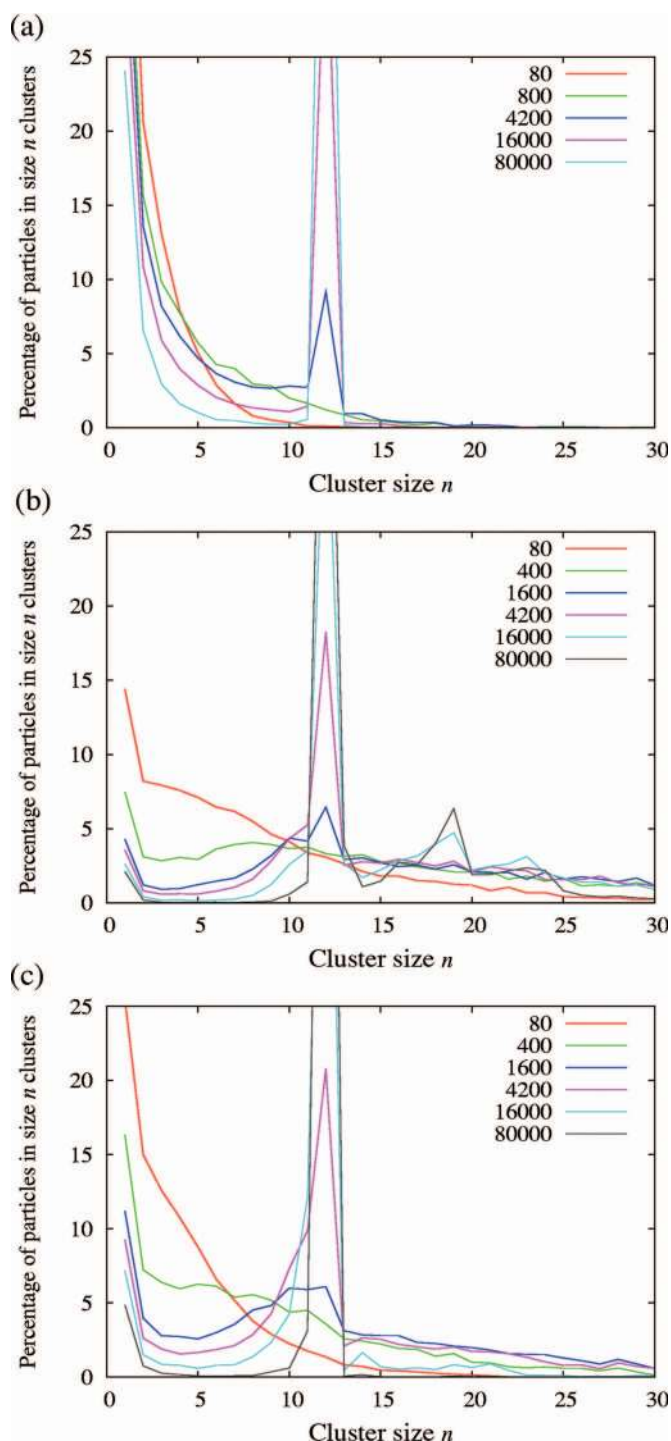

FIG. 7. (Color) The fraction of particles in clusters of a certain size as a function of cluster size at different simulation times for parameters (a) close to the transition to a monomer gas, $T=0.16$, $\sigma=0.4$, (b) close to the transition to a liquid, $T=0.14$, $\sigma=0.55$, and (c) giving optimal conditions for assembly, $T=0.14$, $\sigma=0.45$. In each case the data are an average over 100 simulations, where each simulation is of 1200 particles at a number density of $0.15\sigma_{LJ}^{-3}$. Each line is labeled by the number of MC cycles elapsed.

8(a) shows the yield of icosahedra at a range of particle densities, for the value of σ that we found to be optimal for a density of $0.15\sigma_{LJ}^{-3}$. The dependence of the yield on density is weak, with markedly decreased yields only observed at very low densities, where the times required for the particles to diffuse and reach each other are significantly longer. At very low densities, assembly also becomes comparatively
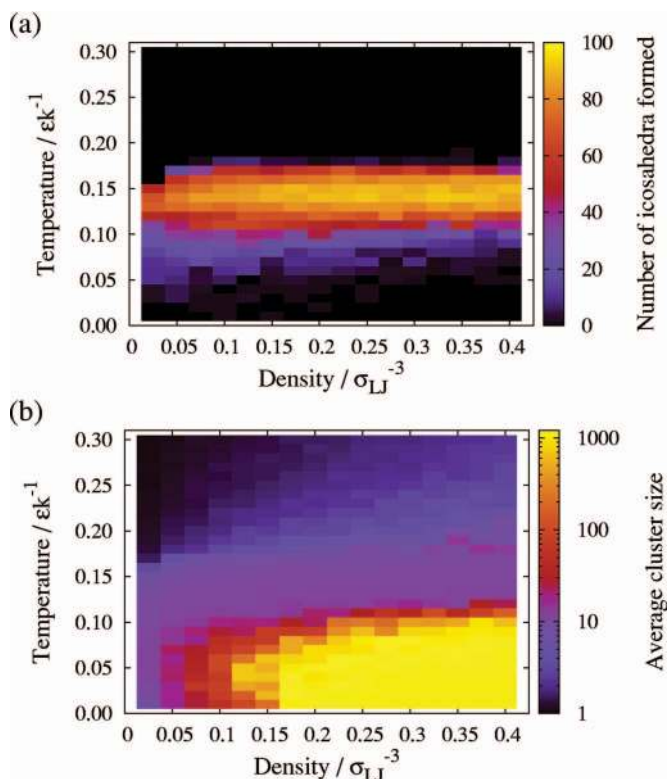
FIG. 8. (Color) (a) The number of icosahedra formed and (b) the mean cluster size (averaged over particles) after 80 000 MC cycles as a function of density and temperature for 1200 particles at a fixed patch width of $\sigma = 0.45$.

more successful at low temperatures. This small effect arises because the formation of large aggregates, which generally competes with the assembly of icosahedra at lower temperature, is disfavored at low density [Fig. 8(b)], and so the successful growth of icosahedra by monomer addition is somewhat enhanced.

F. Effect of inaccurate patch placement

If self-assembling structures are to be constructed from nanoparticles and colloids, experimental methods will have to be developed to pattern the surfaces of the particles appropriately with "sticky" and "nonsticky" regions. Since a certain degree of variability is intrinsic to synthetic colloids, it is useful to consider how robust the self-assembly process is with respect to imperfect placement of the patches. These factors may also be important in the consideration of the evolution of proteins that self-assemble into some quaternary structure, as the robustness of self-assembly will have a strong effect on how easily such proteins can evolve.

In order to examine the effect of imperfect patch placement on assembly, simulations were performed where a random noise term was added to the patch vectors for each particle, with the noise for each patch on each particle being generated separately. The modified patch vectors $\hat{\mathbf{p}}$ are given by
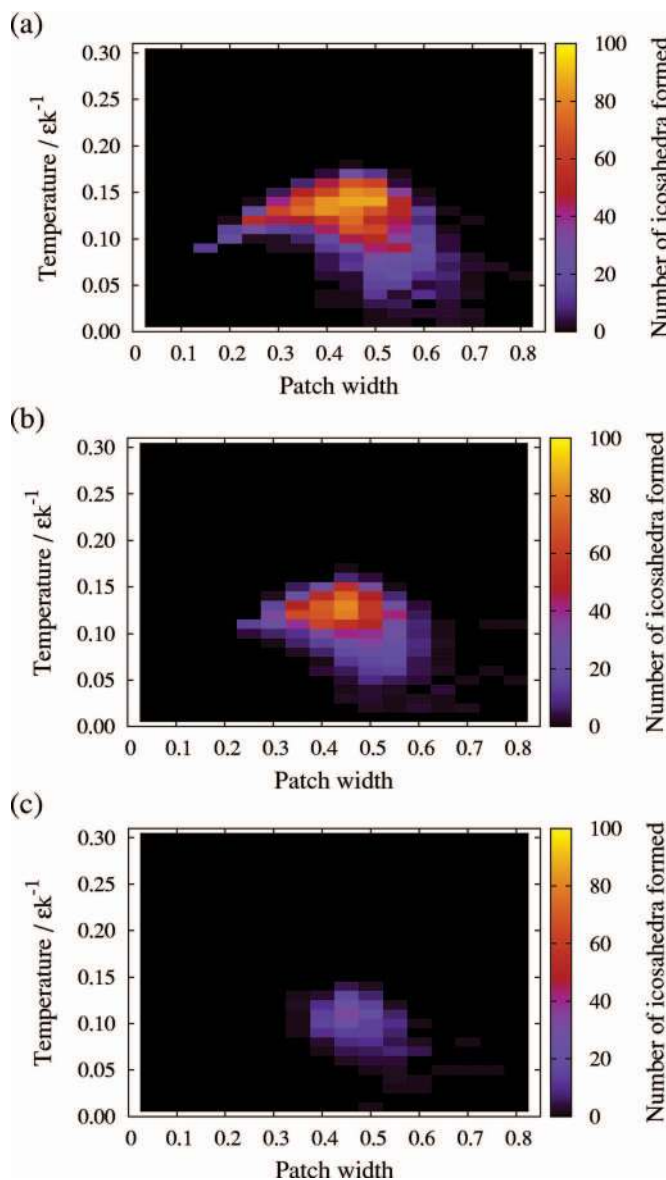

FIG. 9. (Color) The effect of noisy patch placement. The heat maps show the numbers of icosahedra formed after 80 000 MC cycles as a function of the patch width σ (measured in radians) and the temperature for 1200 particles at a number density of $0.15\sigma_{LJ}^{-3}$ for different noise levels in the patch placement: (a) $\mu=0.1$, (b) $\mu=0.2$, and (c) $\mu=0.3$.

$$\hat{\mathbf{p}} = \frac{\hat{\mathbf{p}}_0 + \mu\hat{\mathbf{p}}_1}{|\hat{\mathbf{p}}_0 + \mu\hat{\mathbf{p}}_1|}, \tag{7}$$

where $\hat{\mathbf{p}}_0$ is the original (noiseless) patch vector, $\hat{\mathbf{p}}_1$ is a random unit vector, and μ is a parameter that determines the magnitude of the noise in the patch placement. The average angle of deviation of $\hat{\mathbf{p}}$ from $\hat{\mathbf{p}}_0$ is approximately linear in μ, with a proportionality constant of 0.78 rad.

Analogous plots to Fig. 3(a) are depicted in Fig. 9 for increasing values of the parameter μ. A low level of noise ($\mu=0.1$) seems to have little effect on the formation of icosahedra. Yields away from the optimum parameters are generally slightly depressed and the temperature above which icosahedra are not observed is slightly reduced. With increasing noise levels, however, the region in the (σ, T) plane where assembly is possible shrinks rapidly and by $\mu=0.3$ it
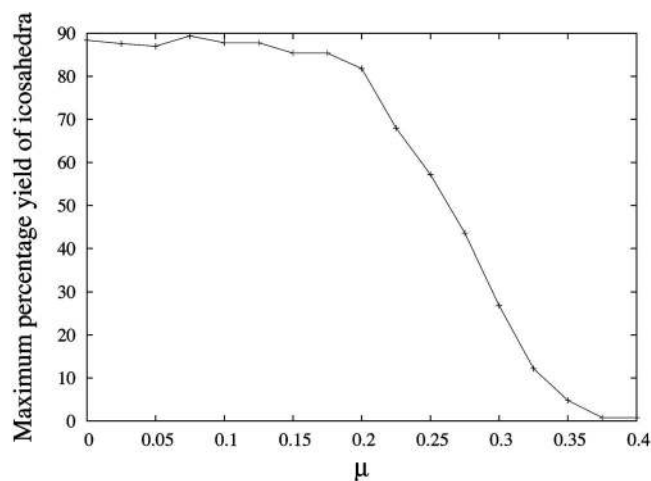
FIG. 10. The maximum yield of icosahedra as a function of the patch placement noise parameter μ. In each case the yield is the average of five simulations at the values of T and σ which give the maximum yield for that value of μ. Each simulation contains 1200 particles, at a number density of $0.15\sigma_{LJ}^{-3}$.
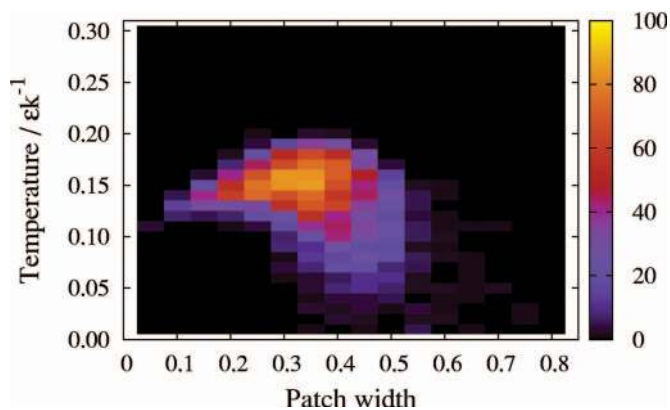

FIG. 11. (Color) Icosahedron yields using ring patches. The heat map shows the number of icosahedra formed after 80 000 MC cycles as a function of the patch width σ (measured in radians) and the temperature for 1200 particles at a number density of $0.15\sigma_{LJ}^{-3}$, using a single ring patch on each particle with $\nu=1.017$ rad.

has nearly disappeared. The noise has the general effect of reducing the binding energy of an icosahedral structure. Thus, the transition temperature to a monomer gas is reduced, as is the patch width above which larger aggregates become competitive. At low σ, i.e., narrow patches, the reduction in the binding energy is more pronounced and icosahedra again become unstable.

The maximum yield of icosahedra as a function of μ is shown in Fig. 10. Interestingly, the maximum yield is virtually independent of noise level up to $\mu=0.2$, after which it falls off rapidly. Beyond $\mu=0.35$ there is virtually no successful assembly. $\mu=0.2$ corresponds to a mean deviation of 0.16 rad, or around 9°, and may represent a target for the accuracy of patch placement in experiments, although one would expect the threshold to be somewhat dependent on the target structure.

G. Use of a single ring patch

The experimental production of five distinct patches that are accurately placed on a nanoparticle or colloidal surface is likely to be very challenging. One potentially easier alternative may be the creation of a continuous attractive ring on the particle's surface.[38] To mimic the effect of the discrete patches the ring should pass through the positions at which the individual patches would otherwise have been placed.

To describe such a case the angular modulation of the interaction potential in the model [Eq. (4)] is replaced by

$$G_{ij}(\hat{\mathbf{r}}_{ij},\Omega_i) = \exp\left[-\frac{(\theta_{k_{ij}}-\nu)^2}{2\sigma^2}\right],\tag{8}$$

where the orientation of the ring patch is described by a patch vector that passes through the center of the ring, ν is the cone angle from the patch vector at which the attraction is greatest, and σ is a measure of the angular width of the ring. The yields of icosahedra for particles with a single ring patch at the appropriate value of ν, namely $\nu=1.017$ rad, are shown in Fig. 11.

A number of comparisons can be made to the simulations where five separate patches were used. First, the maximum yield observed when using a ring patch is the same to within statistical error (85% as opposed to 88% obtained with five discrete patches). It seems that the loss of specificity in changing to a ring patch is not too important, since the value of ν is chosen such that the most favorable configuration is still one where each of the particles has five neighbors. Second, the icosahedra are also stable to a slightly higher temperature when using ring patches, because each particle in an icosahedron is constrained in only two orientational degrees of freedom rather than three, reducing the entropy difference between the icosahedra and the monomeric vapor. Third, the optimal region is shifted to lower σ, because for a given value of σ more of the surface of the particles is attractive than for particles with discrete patches at the same value of σ. Consequently liquidlike aggregates become competitive at lower values of σ. Finally, the assembly process is somewhat less vulnerable to kinetic traps, because the continuous ring patch facilitates rearrangement of the particles.

Our results suggest that ring patches represent an attractive alternative method of patterning particle surfaces to form specific target structures, with prospects for easier synthesis and comparable yields to those obtained using discrete patches.

H. Effect of a short-range potential

All of the previous simulations have used an orientationally modulated version of the Lennard–Jones potential given in Eq. (1). However, the interaction potentials for real colloids and proteins tend to be much shorter in range.[59] Using a longer range potential has computational advantages, but in order to investigate the effect of the range of the potential, a set of simulations was performed using a generalized Lennard–Jones potential with exponents of 20 and 10,

$$V_{LJ}^{20-10}(r) = 4\epsilon\left[\left(\frac{\sigma_{LJ}}{r}\right)^{20} - \left(\frac{\sigma_{LJ}}{r}\right)^{10}\right].\tag{9}$$

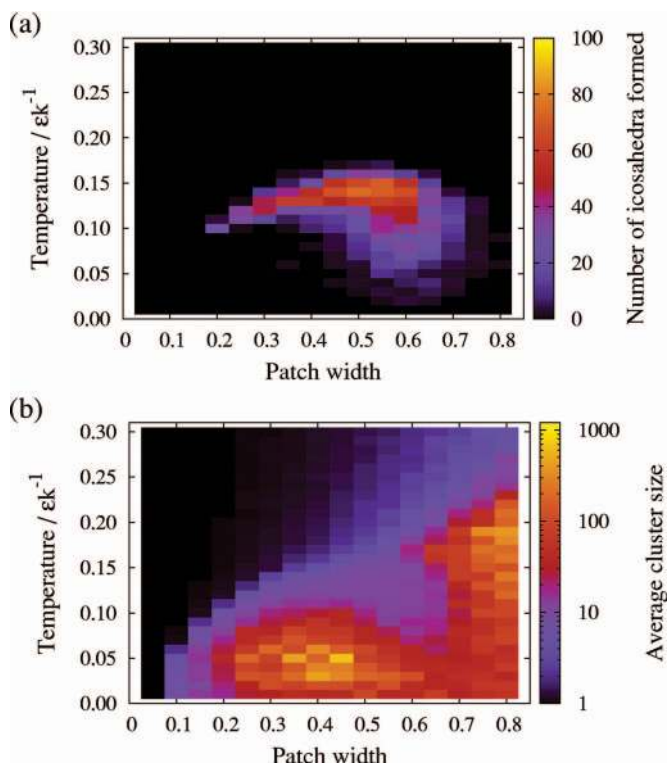The results of these simulations are shown in Fig. 12.

FIG. 12. (Color) (a) The number of icosahedra formed and (b) the mean cluster size (averaged over particles) after 80 000 MC cycles using a shorter-range 20–10 potential as a function of the patch width σ and temperature for 1200 particles at a number density of $0.15\sigma_{LJ}^{-3}$.

semble into monodisperse icosahedral clusters. We found that, given the right conditions, the assembly process was rapid, efficient, and robust. By performing arrays of simulations at different temperatures and using different patch widths, we were able to map out the regions of parameter space in which assembly is successful, and to identify the major competing behaviors arising in the regions of parameter space which did not lead to successful assembly.

From these results we can learn something about the general principles required when designing objects to self-assemble. The optimal patch width represents a compromise between the energetic and kinetic requirements for self-assembly. As the patch width is decreased, the energy gap between the target structure and other possible competing structures increases. By contrast, kinetic accessibility improves as σ increases, because the system is able to escape from incorrect configurations more easily. The lesson is that the interactions need to be specific enough to sufficiently favor the target structure, but that overspecifying them can inhibit the dynamics of assembly.

We also found that the dynamic pathways to self-assembly can be complex and nonintuitive. We observed, for example, that, besides the better known nucleation pathways, clusters could also form through larger disordered intermediates. These mechanisms are less likely to be important in biological systems, but may be relevant for synthetic self-assembling systems. In our simulations the optimal yields were found under conditions where both mechanisms made contributions to the assembly.

Assembly via nucleation pathways appears to be much more successful close to the maximum temperature at which icosahedra are stable. Kinetic traps are avoided and alternative processes such as the formation of liquid droplets are inhibited. Since high temperatures are equivalent to weak bonding interactions, this finding is consistent with the experimental observation that protein-protein interactions in virus capsids are generally weak.[24]

The assembly process is not strongly affected by the particle density. Changing the range of the potential also has little effect on the yields at long times (although assembly mechanisms involving liquid droplets are inhibited by a shorter range potential). The relative independence of the behavior of the system with respect to these factors indicates that our findings are quite general and not just specific to the parameter range on which we focussed.

The self-assembly process is encouragingly robust with respect to the design and placement of patches. Simulations in which the patches were positioned with random inaccuracies still produced almost unchanged yields of correctly formed icosahedra up to a threshold level of average error. While most of our simulations used particles with five discrete patches, we found that particles with a single ring-shaped patch, which may be easier to produce experimentally, assembled with a similar level of efficiency. Both of these findings may be seen as very encouraging for experimentalists seeking to synthesize self-assembling nanoparticles and colloids.

Our study focused on a simple model in order to be able to explore the design space in detail. Although the patchy

The overall trends are similar to those for the longer range potential. However, yields are generally reduced, because the shorter range potential makes it less likely for particles to enter each other's attractive potential wells and form bonds, and because the barriers to rearrangements increase.[60] As a result the entire self-assembly process proceeds more slowly. At longer simulation times the yields recover, with a maximum yield of icosahedra of 97% after 1.2×10^6 MC cycles. Another general effect is that because of the slower bond formation, any aggregates that form will tend to be reduced in size [Fig. 12(b)].

The shorter range of the 20–10 potential energetically destabilizes the liquid phase,[61,62] thus increasing the maximum value of σ at which the system is stable as a collection of icosahedra. Hence, icosahedra are able to assemble up to higher σ. Furthermore, since the budding mechanism of assembly requires the formation of a liquid droplet, assembly by nucleation is also comparatively more dominant than in the case of the 12–6 potential.

Decreasing the range of the potential also has some other small effects, as seen in Fig. 12(a). The vibrational entropy of the icosahedra is reduced, lowering T_{clust} slightly. There is also a small increase in the value of σ below which the kinetic suppression of icosahedron formation occurs, since at low σ the short range further increases the difficulty of assembly.

IV. CONCLUSIONS

By making use of a minimal model, we have been able to simulate the behavior of patchy particles designed to as-

spheres of our model clearly lack the complexity of the units involved in monodisperse self-assembly in biological systems, the comprehensive understanding that this simplicity allows us to achieve should provide useful insights into biological examples. Furthermore, this feature may be an advantage when considering the design of synthetic self-assembling building blocks.

The target structure used in this work was relatively simple and formation of monodisperse clusters was correspondingly rapid. In future work we intend to explore how the ease of assembly depends on the geometry of the target structure and to investigate the limits on the complexity of structures that can be generated from this relatively simple interaction scheme. In addition, there are various features which could be added to the potential, most notably a dependence on torsional angles. We expect that a torsional dependence would greatly increase the range of achievable target structures and also produce behavior more closely mirroring that of biological systems.

ACKNOWLEDGMENTS

[1] G. M. Whitesides and B. Grzybowski, Science **295**, 2418 (2002).
[2] J. B. Bancroft, G. J. Hills, and R. Markham, Virology **31**, 354 (1967).
[3] D. M. Salunke, D. L. D. Caspar, and R. L. Garcea, Biophys. J. **56**, 887 (1989).
[4] B. Rombaut, R. Vrijsen, and A. Boeyé, Virology **177**, 411 (1990).
[5] P. E. Prevelige, D. Thomas, and J. King, Biophys. J. **64**, 824 (1993).
[6] D. L. D. Caspar and A. Klug, Cold Spring Harbor Symp. Quant. Biol. **27**, 1 (1962).
[7] R. Zandi, D. Reguera, R. Bruinsma, W. M. Gelbart, and J. Rudnick, Proc. Natl. Acad. Sci. U.S.A. **101**, 15556 (2004).
[8] T. Chen, Z. Zhang, and S. C. Glotzer, Proc. Natl. Acad. Sci. U.S.A. **104**, 717 (2007).
[9] A. Zlotnick, J. Mol. Recognit. **18**, 479 (2005).
[10] D. Endres, M. Miyahara, P. Moisant, and A. Zlotnick, Protein Sci. **14**, 1518 (2005).
[11] R. Zandi, P. van der Schoot, D. Reguera, W. Kegel, and H. Reiss, Biophys. J. **90**, 1939 (2006).
[12] M. Hemberg, S. N. Yaliraki, and M. Barahona, Biophys. J. **90**, 3029 (2006).
[13] S. D. Hicks and C. L. Henley, Phys. Rev. E **74**, 031912 (2006).
[14] D. C. Rapaport, Phys. Rev. E **70**, 051905 (2004).
[15] M. F. Hagan and D. Chandler, Biophys. J. **91**, 42 (2006).
[16] D. Endres and A. Zlotnick, Biophys. J. **83**, 1217 (2002).
[17] J. M. Johnson, J. Tang, Y. Nyame, D. Willits, M. J. Young, and A. Zlotnick, Nano Lett. **5**, 765 (2005).
[18] H. D. Nguyen, V. S. Reddy, and C. L. Brooks III, Nano Lett. **7**, 338 (2007).
[19] T. Zhang and R. Schwartz, Biophys. J. **90**, 57 (2006).
[20] A. Zlotnick, J. Mol. Biol. **366**, 14 (2007).
[21] K. N. Parent, S. M. Doyle, E. Anderson, and C. M. Teschke, Virology **340**, 33 (2005).
[22] K. N. Parent, A. Zlotnick, and C. M. Teschke, J. Mol. Biol. **359**, 1097 (2006).
[23] K. N. Parent, M. M. Suhanovsky, and C. M. Teschke, J. Mol. Biol. **365**, 513 (2007).
[24] A. Zlotnick, Virology **315**, 269 (2003).
[25] J. M. Johnson, D. A. Willits, M. J. Young, and A. Zlotnick, J. Mol. Biol. **335**, 455 (2004).
[26] G. L. Casini, D. Graham, D. Heine, R. L. Garcea, and D. T. Wu, Virology **325**, 320 (2004).
[27] C. M. Teschke, A. McGough, and P. A. Thuman-Commike, Biophys. J. **84**, 2585 (2003).
[28] R. P. Goodman, I. A. T. Schaap, C. F. Tardin, C. M. Erben, R. M. Berry, C. F. Schmidt, and A. J. Turberfield, Nature (London) **310**, 1661 (2005).
[29] A. van Blaaderen, Nature (London) **439**, 545 (2006).
[30] Z. Li, D. Lee, M. F. Rubner, and R. E. Cohen, Macromolecules **38**, 7876 (2005).
[31] Y.-S. Cho, G.-R. Yi, J.-M. Lim, S.-H. Kim, V. N. Manoharan, D. J. Pine, and S.-M. Yang, J. Am. Chem. Soc. **127**, 15968 (2005).
[32] K.-H. Roh, D. C. Martin, and J. Lahann, Nat. Mater. **4**, 759 (2005).
[33] C. E. Snyder, A. M. Yake, J. D. Feick, and D. Velegol, Langmuir **21**, 4813 (2005).
[34] A. M. Jackson, J. W. Myerson, and F. Stellacci, Nat. Mater. **3**, 330 (2004).
[35] G. A. DeVries, M. Brunnbauer, Y. Hu, A. M. Jackson, B. Long, B. T. Neltner, O. Uzun, B. H. Wunsch, and F. Stellacci, Science **315**, 358 (2007).
[36] R. Lévy, Z. Wang, L. Duchesne, R. C. Doty, A. I. Cooper, M. Brust, and D. G. Fernig, ChemBioChem **7**, 592 (2006).
[37] K. Van Workum and J. F. Douglas, Phys. Rev. E **73**, 031502 (2006).
[38] Z. Zhang and S. C. Glotzer, Nano Lett. **4**, 1407 (2004).
[39] S. C. Glotzer, Science **306**, 419 (2004).
[40] 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, Phys. Chem. Chem. Phys. **9**, 2197 (2007).
[41] E. G. Noya, C. Vega, J. P. K. Doye, and A. A. Louis, J. Chem. Phys. **127**, 054501 (2007).
[42] D. Frenkel and B. Smit, *Computer Simulation of Liquids* (Academic, London, 2001), pp. 48–49.
[43] K. Kikuchi, M. Yoshida, T. Maekawa, and H. Watanabe, Chem. Phys. Lett. **185**, 335 (1991).
[44] G. Tiana, L. Sutto, and R. A. Broglia, Physica A **380**, 241 (2007).
[45] L. Berthier and W. Kob, J. Phys.: Condens. Matter **19**, 205130 (2007).
[46] G. M. Torrie and J. P. Valleau, J. Comput. Phys. **23**, 187 (1977).
[47] D. Frenkel and B. Smit, *Computer Simulation of Liquids* (Academic, London, 2001), pp. 192–196.
[48] D. J. Wales and J. P. K. Doye, J. Phys. Chem. A **101**, 5111 (1997).
[49] D. J. Wales and H. A. Scheraga, Science **285**, 1368 (1999).
[50] J. A. Northby, J. Chem. Phys. **87**, 6166 (1987).
[51] D. J. Wales, Philos. Trans. R. Soc. London, Ser. A **363**, 357 (2005).
[52] J. P. K. Doye, M. A. Miller, and D. J. Wales, J. Chem. Phys. **111**, 8417 (1999).
[53] J. D. Bryngelson, J. N. Onuchic, N. D. Socci, and P. G. Wolynes, Proteins **21**, 167 (1995).
[54] P. R ten Wolde and D. Frenkel, Phys. Chem. Chem. Phys. **1**, 2191 (1999).
[55] P. R. ten Wolde and D. Frenkel, Science **277**, 1975 (1997).
[56] D. Vivarès, E. W. Kaler, and A. M. Lenhoff, Acta Crystallogr., Sect. D: Biol. Crystallogr. **61**, 819 (2005).
[57] Z. Tavassoli and R. P. Sear, J. Chem. Phys. **116**, 5066 (2002).
[58] A. Zlotnick and S. J. Stray, Trends Biotechnol. **21**, 536 (2003).
[59] D. Rosenbaum, P. C. Zamora, and C. F. Zukoski, Phys. Rev. Lett. **76**, 150 (1996).
[60] M. A. Miller, J. P. K. Doye, and D. J. Wales, J. Chem. Phys. **110**, 328 (1999).
[61] M. H. J. Hagen and D. Frenkel, J. Chem. Phys. **101**, 4093 (1994).
[62] J. P. K. Doye and D. J. Wales, Science **271**, 484 (1996).