THE JOURNAL OF CHEMICAL PHYSICS **135**, 104115 (2011)

Mechanisms of kinetic trapping in self-assembly and phase transformation

Michael F. Hagan,[1] Oren M. Elrad,[1] and Robert L. Jack[2,a)]
[1]*Department of Physics, Brandeis University, Waltham, Massachusetts 02254, USA*
[2]*Department of Physics, University of Bath, Bath BA2 7AY, United Kingdom*

(Received 6 May 2011; accepted 18 August 2011; published online 14 September 2011)

In self-assembly processes, kinetic trapping effects often hinder the formation of thermodynamically stable ordered states. In a model of viral capsid assembly and in the phase transformation of a lattice gas, we show how simulations in a self-assembling steady state can be used to identify two distinct mechanisms of kinetic trapping. We argue that one of these mechanisms can be adequately captured by kinetic rate equations, while the other involves a breakdown of theories that rely on cluster size as a reaction coordinate. We discuss how these observations might be useful in designing and optimising self-assembly reactions. © *2011 American Institute of Physics*. [doi:10.1063/1.3635775]

I. INTRODUCTION

In self-assembly processes, simple components come together spontaneously to form ordered products. Such processes abound in biology, where the ordered structures might be the outer shells of viruses,[1–5] extended one-dimensional filaments that make up the cytoskeleton,[6] or ordered arrays of proteins on the surface of bacteria.[7] In other areas of nanoscience, self-assembled nanostructures made from customised DNA oligomers are being used to build ever more complex structures,[8] and the possibility of tailoring interactions between colloidal particles to assemble diverse ordered structures and phases is also an area of active research.[9]

Here, we concentrate on self-assembly processes that occur without energy input. That is, we consider systems of interacting components, relaxing towards ordered equilibrium states. In order to assemble a particular ordered structure, there are two requirements that must be met. Firstly, interactions between particles must be chosen so that the ordered state minimises the system's free energy and is, therefore, stable at equilibrium. Secondly, one must address dynamic questions: how long does it take for a system to reach its ordered equilibrium state, and how can interparticle interactions be optimised to facilitate rapid and effective assembly?

Even if the equilibrium state of a system is ordered, there are many scenarios in which formation of the relevant order occurs extremely slowly. In self-assembly, this is known as "kinetic trapping." In recent years, several studies[4,5,10–18] have observed that self-assembly is most efficient when structures are stabilised by large numbers of relatively weak interactions. In particular, while strong interparticle bonds stabilise the ordered equilibrium state, they are also associated with kinetic trapping effects that frustrate the assembly process. Instead, effective self-assembly proceeds by relatively transient bond formation, including frequent bond-breaking events.

Long ago, Caspar and Klug[1] drew an analogy between viral capsid assembly and phase changes such as crystallisation. Indeed, strong bonds may act to frustrate crystallisation dynamics and other phase change processes, just as they do in self-assembly. For example, phase transformations may be unobservable due to the formation of amorphous aggregates,[19] or to gelation.[20] These are also types of kinetic trapping, in that these systems are prevented from equilibrating into their ordered free-energy minima.

Our aim in this article is to consider mechanisms for kinetic trapping that are generic, in that they apply across a large range of systems. To this end, we consider two processes that are apparently quite different: self-assembly of model viral capsids and phase separation in the two-dimensional Ising lattice gas. Viral capsids are monodisperse closed shells made of many identical smaller particles,[1] while the ordered structures that form in the lattice gas model are extended two-dimensional clusters. Yet, despite these differences between their assembled products, we find that these models exhibit some striking similarities and common features—we interpret these similarities as evidence for generic mechanisms by which dynamical effects frustrate the formation of ordered states. Both crystallisation and capsid formation involve particles undergoing rearrangement toward an ordered free energy minimum, so we use "self-assembly" as a broad term which covers both these processes.

In contrast to the studies of kinetic trapping based on energy landscapes[21–23] or schematic models of particle aggregation,[19,24] we do not attempt a detailed analysis of the structures of the disordered states that lead to kinetic trapping in our model systems: at that level, the capsid and lattice gas models are very different. Rather, our aims are firstly to identify common features at the level of the assembly dynamics, and secondly to understand how the presence of kinetic traps may act to frustrate self-assembly and phase change, independent of the structural features of the disordered states. In particular, we identify two generic trapping mechanisms, one of which can be captured by "classical" theories of phase change,[25,26] while the other is associated with a breakdown of these theories. We discuss how deviations from the classical theory might be identified and characterised, based on ideas proposed by some of us.[5]

[a)] Author to whom correspondence should be addressed. Electronic mail: r.jack@bath.ac.uk.

II. MODELS

A. Viral capsid model

We first describe a model for the self-assembly of empty icosahedral viral capsids. The model represents capsid proteins as rigid bodies ("subunits") with excluded volume geometries and orientation-dependent interactions. The lowest energy structure is an icosahedral shell consisting of 20 subunits (details are given in Appendix A and Fig. 8 as well as in Ref. 27). This model was used to examine the assembly of icosahedral viruses around a polymer in Ref. 27 and is similar to models used by Rapaport *et al.*[12,28] and Nguyen *et al.*[15] in simulations of empty capsid assembly. Each subunit could correspond to a "capsomer" comprising a trimer of proteins that form a $T = 1$ capsid.

The binding energy associated with each interaction site on a subunit is $\varepsilon_{\rm b}$ and we take Boltzmann's constant $k_{\rm B} = 1$, so that the relevant dimensionless parameter is $\varepsilon_{\rm b}/T$. When bonds are strong, the equilibrium state of the system has nearly all particles in complete capsids; when bonds are weak, most particles are isolated monomers. These regimes are separated by a crossover which is characterised by the "critical capsomer concentration" $\rho_{\rm cc}$: this is the number density of subunits for which the equilibrium state contains half of the particles in complete capsids and is analogous to the "critical micelle concentration" in the theory of surfactants. In Appendix C, we derive an approximate formula for $\rho_{\rm cc}$ and its dependence on $\varepsilon_{\rm b}$: we estimate energetic and entropic contributions to the capsid free energy using a combination of umbrella sampling and counting arguments. The result is shown in Fig. 1(a), together with an indication of the state points considered in this work.

To simulate the dynamical process of self-assembly, we use over-damped Brownian dynamics for the capsid subunits as in Ref. 27 using periodic boundary conditions and a second order predictor-corrector algorithm.[29] The capsomer subunits have anisotropic translational and rotational diffusion constants calculated using HYDROSUB7.C.[30] To obtain dimensionless units, we rescale lengths by $\sigma_{\rm b}$, which is the diameter of one of the spheres that comprise the excluded volume of the capsomer; times are measured in units of t_0, which is the Brownian time for a single such sphere. Further details of the model are given in Appendix A.

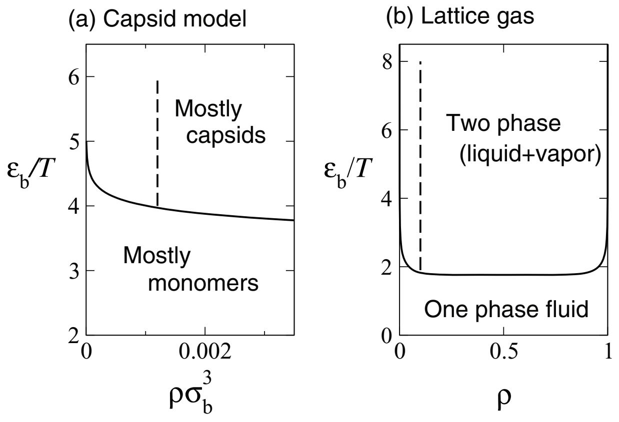

FIG. 1. Thermodynamic behaviour of the capsid and lattice gas models, with dashed lines indicating the state points considered in this paper. (a) Capsid model: The solid line is an approximation to the "critical capsomer concentration" $\rho_{\rm cc}$, where the equilibrium state of the model has half of the particles in complete capsds. (b) Lattice gas model: The solid line is the binodal, which separates regimes of one-phase and two-phase behaviour.

B. Viral capsid assembly in the canonical ensemble.

In Fig. 2(a), we show the results from simulations of self-assembly at constant particle number, volume, and temperature (NVT) for various values of the interaction energy $\varepsilon_{\rm b}$. The initial conditions for the simulations have subunits with random positions and orientations. We measure the number of perfect capsids in the system (a perfect capsid is defined as a cluster with exactly 20 subunits, each of which has its maximum number of three bonds). We associate the fraction of capsomer subunits in perfect capsids $n_{\rm capsid}(t)$ with the *yield* of the assembly process.[31]

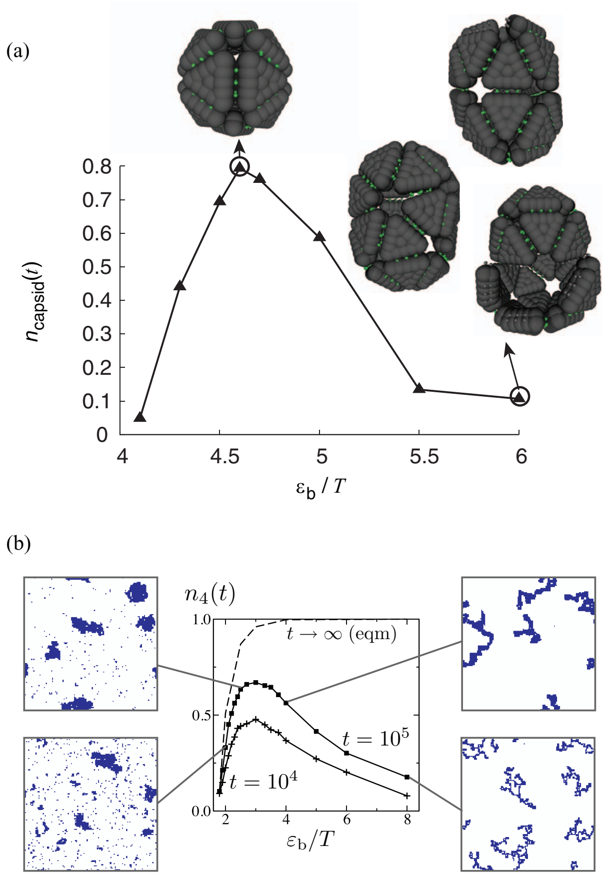

FIG. 2. Assembly from a disordered state. (a) Dynamical capsid assembly yields in the *NVT* ensemble. The fraction of subunits in well-formed capsids, $n_{\rm capsid}(t)$ is shown for $t = 210\,000 t_0$ as a function of the binding interaction parameter $\varepsilon_{\rm b}$. Snapshots exemplify typical clusters at the circled points. Green attractor pseudoatoms are experiencing favorable interactions, while gray attractors are not. The size of the attractors indicates the length scale of their interaction. The system contains $N = 500$ trimer subunits in a box of sidelength $L = 74\sigma_{\rm b}$. (b) Phase change in the lattice gas at density $\rho = 0.1$. The binodal is located at $\varepsilon_{\rm b}/T = 1.86$. The assembly yield is the fraction of particles that have four bonds $n_4(t)$. The snapshots show representative configurations at time $t = 10^5$. The dashed line shows the yield that would be obtained by equilibrating a very large system: this "thermodynamic yield" is monotonic in the bond strength $\varepsilon_{\rm b}$ while the yield at fixed time is non-monotonic. The lattice size is $L = 128$.

As anticipated in Sec. I, the yield depends on a combination of thermodynamic and kinetic effects. For weak bonds (small ε_b/T), there is little thermodynamic drive to assemble and no capsids are formed. For strong bonds (large ε_b/T), the thermodynamic drive to assembly is strong, but the system is vulnerable to kinetic trapping and forms disordered clusters of subunits instead of perfect capsids. Optimal assembly takes place in an intermediate range $\varepsilon_b/T \approx 4.5$.[32] In later sections, we will analyse the interplay of thermodynamic and kinetic effects in more detail.

C. Lattice gas model of self-assembly

We also consider 'self-assembly' in an Ising lattice gas containing N particles on a (two-dimensional) square lattice with $V = L^2$ sites. Particles may not overlap, so the occupancy of site i is $n_i \in \{0, 1\}$. Particles on neighbouring sites form bonds with energy ε_b so that the energy of a configuration is

$$E = -\varepsilon_b \sum_{\langle ij\rangle} n_j n_j, \tag{1}$$

where the sum runs over (distinct) pairs of nearest neighbours. Working with a fixed number of particles, the system is unstable to phase separation at low temperatures, forming dense (liquid) and dilute (gas) phases. To make an analogy with self-assembly, we start with the particles in a disordered configuration, and measure the rate with which order is formed (see also Ref. 14). To quantify the yield of the assembly process, we measure the number of particles that have bonds to all four of their neighboring sites: we denote this number by N_4 and we write $n_4(t) = (1/N)\langle N_4(t)\rangle$ for the fraction of particles with four bonds. The model evolves in time according to a Monte Carlo (MC) procedure that involves cluster moves, chosen to produce dynamically realistic trajectories, at least qualitatively.[7,33] The method that we use is close to that described in Ref. 34. In each MC move, we select a seed particle and use it to build a cluster as follows. For each particle bonded to the seed particle, we conduct a Monte Carlo trial, adding it to the cluster with probability $p_c = 1 - e^{-\varepsilon_b/T}$. This process is then repeated recursively: for those particles that have been added, we use the same MC trial to decide whether particles bonded to them are added in turn. Taking the resulting cluster, we propose a move in a random direction. This move is rejected if this proposed move would lead to more than one particle on any site. Otherwise, the move is accepted with a probability $p_a = 1/n^2$, where n is the size of the cluster to be moved. A MC sweep consists of N moves and the time is measured in MC sweeps. The choices of p_c and p_a ensure that the procedure obeys detailed balance with respect to a Boltzmann distribution whose energy is given by Eq. (1), and also that large clusters of particles diffuse freely through the system with a diffusion constant consistent with Brownian dynamics $D(n) \propto 1/n$. This dynamical scheme represents a schematic description of particles with short-ranged attractions moving through a solvent.[35]

The relevant variables in our NVT simulations of this system are the dimensionless bond strength ε_b/T and density $\rho = N/V$. The phase behaviour as a function of these two parameters is well-known: the system will eventually phase separate at all temperatures below the binodal (i.e., when $\sinh^4(\varepsilon_b/2T) > 1/[1-(2\rho-1)^8]$, see Ref. 36). The location of the binodal is indicated in Fig. 1(b). For the values of ε_b/T that we consider, phase separation does not take place by a single rare nucleation event but rather by multiple clusters that form quickly in the system and then grow. Thus, we are in the regime of spinodal decomposition, at least in the sense that there is no significant free energy barrier to nucleation.

To obtain the dynamical measurement of assembly yield shown in Fig. 2(b), we initialise the particles in random positions and propagate the dynamics. The particles assemble into clusters: for temperatures below the binodal, these clusters will grow until their size becomes comparable to the whole system. However, for the times we consider, domains are much smaller than the system size so the system is always far from equilibrium. (This also means that finite size effects are small in our simulations over the parameter range shown.) As in the viral capsid model, the yield $n_4(t)$ is non-monotonic in the bond strength ε_b/T: the thermodynamic driving force to assemble is small when ε_b/T is small, while kinetic trapping sets in for large ε_b/T.

III. STEADY STATE ENSEMBLE: RATE AND QUALITY OF ASSEMBLY

A. Steady state ensemble

We now discuss a self-assembling steady state, using ideas that have been exploited in the studies of nucleation and phase transformation.[37] Given a self-assembling model system such as the capsid or lattice gas model, we simulate the dynamics of the system in the usual way, except that we periodically remove large clusters of particles (subunits) from the system. We refer to clusters removed in this way as the *products* of the self-assembly process. The morphologies of the product clusters are stored for later analysis, and we then reintroduce free particles into the system at random positions so that the total number of particles in the system remains constant. To make connection with experiment, we imagine a continuous assembly process, where free particles (subunits) are fed into the system and large assembled products are removed, perhaps by exploiting their tendency to sediment. On starting the system in an initially random configuration, it settles down into a time-translationally invariant steady state in which product clusters are continuously assembling. This feature of the steady state ensemble allows time-invariant averages to be taken during productive assembly in contrast to the NVT quenches. The criteria for identifying large clusters depend on the model being simulated and are described below.

We show in Figs. 3 and 4 that the physical processes taking place in NVT simulations and in the steady state are similar. That is, effective assembly and kinetic trapping take place at similar system parameters, and the structures of disordered kinetically trapped clusters are also similar between ensembles. Specifically for the capsid system, we find that the structures of growing clusters in the steady state ensemble closely approximate those found in NVT simulations at the end of their lag phase, which corresponds to the point at which the first capsids are completing assembly[38] and the

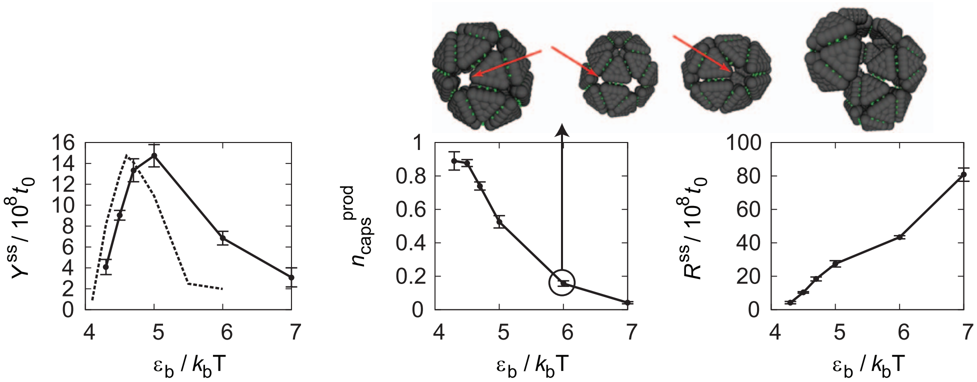

FIG. 3. Steady state ensemble results for the capsid model. We show the yield Y^{ss}, the product "quality" $n_{\mathrm{caps}}^{\mathrm{prod}}$ and the production rate R^{ss}. Four snapshots corresponding to product clusters from the circled perimeter set, $\varepsilon_{\mathrm{b}} = 6.0$ are shown on top of the plots. The steady state simulations had $N = 1000$ trimer subunits in a box with sidelength $L = 93.23\sigma_{\mathrm{b}}$. The red arrows indicate the location of hexameric defects discussed in Sec. V C.

intermediate-sized clusters have built up an "assembly line."[39] More generally, we find similar qualitative behaviour in steady state and NVT simulations, indicating that two ensembles are equally valid for studies of self-assembly in these models. We introduce the steady state ensemble both for computational convenience and to aid the interpretation of our results. Since this ensemble is time-translationally invariant, measurements made at any time during the simulation are statistically equivalent, which facilitates the evaluation of averages and correlation functions during assembly. Also, measurements from a single time contain information about the entire assembly pathway, including all intermediates that are formed between single monomers and final products. We show below that this feature is useful in interpreting data for the dynamical processes considered here, and we believe it has potential application to other self-assembly processes as well.

We use the notation $\langle \cdot \rangle_{\mathrm{ss}}$ for averages within the steady state. For example, the average lattice gas energy $\langle E(t) \rangle_{\mathrm{ss}}$ is obtained by averaging the energy defined in Eq. (1) at a time t in the steady state regime. We also take averages over the product clusters that are formed in the steady state. To be precise, after a simulation has been in the steady state for a time t_{obs}, let the number of product clusters formed in that time be $\mathcal{M}$. Averaging over many such runs, we obtain the rate of product formation, per unit time and unit volume

$$R^{\mathrm{ss}} \equiv \frac{1}{V t_{\mathrm{obs}}} \langle \mathcal{M} \rangle. \qquad (2)$$

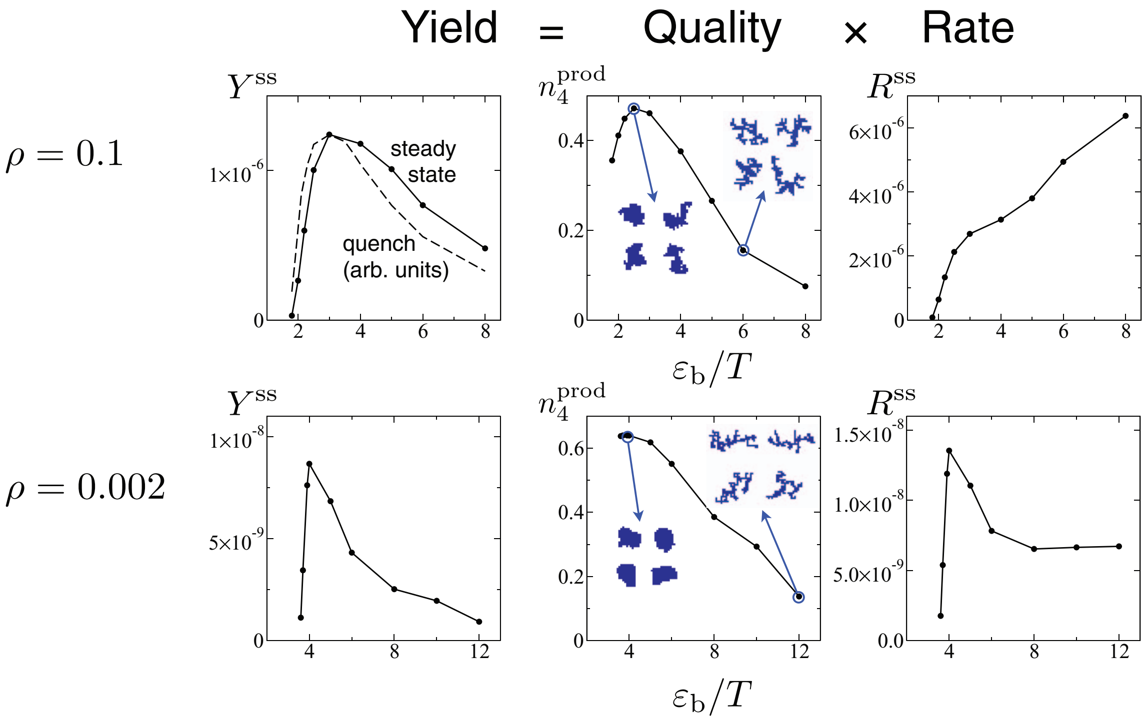

FIG. 4. Steady state ensemble in the lattice gas with $n_{\max} = 100$. At two different densities, we show the yield Y^{ss}, the product "quality" n_4^{prod} and the "production rate" R^{ss}. At $\rho = 0.1$, we compare with the yield $\langle n_4(t) \rangle$ at a time $t = 10^5$ after a "quench" from a disordered state (data from Fig. 2, rescaled for comparison). In the central panels, we show example "product" clusters, to indicate their morphologies.

Labelling the product clusters from a given run by an index $\mu = 1, 2, \ldots, \mathcal{M}$, we may then calculate averages over these product clusters, which we denote by $\langle \cdot \rangle_{\text{prod}}$. For example, in the lattice gas, if $N(\mu)$ is the number of particles in cluster μ then $\langle N(\mu) \rangle_{\text{prod}}$ is obtained by averaging this number over all product clusters.

To make a connection between the steady state ensemble and the NVT simulations of Fig. 2, it is useful to define the steady state yield

$$Y^{\text{ss}} \equiv Q^{\text{prod}} \times R^{\text{ss}}, \tag{3}$$

where Q^{prod} is a measure of the "assembly quality" of the products. For example, in the viral capsid model Q^{prod} is the fraction of product clusters that are perfect capsids, so that the steady state yield Y^{ss} is the production rate of perfect capsids.

B. Viral capsid model in the steady state ensemble

Results from the steady state ensemble of the viral capsid model are shown in Fig. 3. In this model, we define particles to be *connected* if they enjoy a bond with interaction strength $U < -5k_{\text{B}}T$. A cluster of connected particles is identified as a *product* if (i) the cluster has either more than 22 particles or it is a perfect capsid, and (ii) the cluster has remained bonded with at least 17 of the same subunits for a time $t \geq 13.3t_0$. We have in mind that the products of the assembly process be stable long-lived clusters and this second condition limits erroneous identification of weakly bonded short-lived clusters as products.

In Fig. 3, we show the production rate R^{ss}, the yield Y^{ss}, and the product quality Q^{prod}, which is equal to the fraction of product clusters that are perfect capsids $n_{\text{caps}}^{\text{prod}}$. As for the NVT simulations of Fig. 2, we observe that the yield is non-monotonic with respect to binding energy, with an optimum at $\varepsilon_{\text{b}}/T \approx 5$. The similar position of the optimal yield in NVT and steady state simulations indicates that both ensembles are capturing the same physical processes. However, quantitative comparison between yield from the two ensembles is hindered by the fact that Y^{ss} is a rate while $n_{\text{capsid}}(t)$ is a dimensionless number that depends on the measurement time t.

The origin of the optimum in Y^{ss} is a competition between a rate R^{ss} that increases on increasing ε_{b}, and a quality factor Q^{prod} that decreases. (The total production rate increases with ε_{b} over the whole range considered, although it eventually decreases at much higher ε_{b}, for reasons, discussed in Sec. V.)

In terms of kinetic trapping, we find that for large ε_{b}/T product clusters are being formed quickly, but these clusters are of low quality. In later sections, we contrast this scenario with the "stalling" or "starvation" scenarios discussed by Zlotnick and Endres[39] in the context of viral capsid assembly. In that case, kinetic trapping appears as a rate R^{ss} that decreases sharply as ε_{b}/T is large (see also Sec. V). Figure 3 shows that this is not the case for the steady state ensemble with the parameters that we simulate here. We note here that the relative importance of different kinetic trapping mechanisms at a given set of parameter values may differ between the steady state and NVT ensembles, since the recycling of subunits from product clusters in the steady state ensemble provides a continuous supply of monomers that reduces the starvation trap. However, we show below that both the "low quality" and starvation mechanisms may be observed in both ensembles: this reinforces our conclusion that the steady state ensemble is appropriate for studying both trapping mechanisms.

C. Lattice gas model in the steady state ensemble

In the lattice gas model, clusters are identified as products if their size is larger than a maximal cluster size n_{max}. We choose $n_{\text{max}} = 100$ although our results depend only weakly on n_{max}. For the product quality Q^{prod}, we take the fraction of product particles that have four bonds, calculated as

$$Q^{\text{prod}} = n_4^{\text{prod}} = \frac{\langle N_4(\mu) \rangle_{\text{prod}}}{\langle N(\mu) \rangle_{\text{prod}}}, \tag{4}$$

where $N_4(\mu)$ is the number of four-bonded particles in product cluster μ.

Results from the steady state ensemble are shown in Fig. 4 for two different densities ρ. The non-monotonic behaviour of the yield Y^{ss} mirrors the behaviour of the yield $n_4(t)$ found in Fig. 2.[40] At $\rho = 0.1$ (top panels of Fig. 4), the results are similar to those shown for the capsid system in Fig. 3: on increasing ε_{b}/T, the non-monotonic yield arises from a competition between an increasing production rate R^{ss} and a decreasing quality Q^{prod}. However, in simulations at a lower density $\rho = 0.002$, the rate R^{ss} is itself non-monotonic. The scenario that occurs at low densities is consistent with the stalling (monomer starvation) effect,[39] where the system is depleted of free particles, leading to slow cluster growth. But it is the relatively high density ($\rho = 0.1$) scenario in the lattice gas that mimics the data for the viral capsid model shown in Fig. 3. As in that case, kinetic trapping occurs not just because of depletion of free particles, but rather from disordered large clusters or aggregates, examples of which are shown in Fig. 4.

IV. MEASURES OF CLUSTER EQUILIBRATION

We now discuss the relation between cluster quality Q^{prod} and a condition that we call "cluster equilibration." Our idea is that one type of kinetic trapping arises from the disordered aggregates such as those discussed above, and that the importance of these aggregates may be measured through deviations from cluster equilibration.[41]

For a general definition of cluster equilibration, we characterise clusters of particles by their size n and by a second index $\alpha, \beta, \gamma, \ldots$ that indicates their morphology. If $\mathcal{N}_{n,\alpha}$ is the number of clusters with size n and morphology α then our cluster equilibration condition is

$$\frac{\langle \mathcal{N}_{n,\alpha} \rangle}{\langle \mathcal{N}_{n,\gamma} \rangle} = \mathrm{e}^{-(E_{n,\alpha} - E_{n,\gamma})/T}, \tag{5}$$

where $E_{n,\alpha}$ is the energy of a cluster of n particles and morphology α, and the averages might be taken at a fixed time during assembly in the NVT ensemble, or in the steady state ensemble. In words, Eq. (5) states that: "for clusters of size n, the probabilities of different morphologies are Boltzmann-distributed." It seems natural to interpret this as a cluster equilibration condition. (If the clusters have different excluded volumes one might take this into account by replacing the energy with a suitable enthalpy, and any internal entropy of the cluster can also be incorporated through a cluster free energy.) In theoretical treatments of self-assembly based on rate equations or field-theoretic arguments, it is natural to assume that Eq. (5) holds (see Sec. V). We now show that deviations from Eq. (5) are significant throughout the regimes where kinetic trapping is important indicating that such deviations must be taken into account in theories of self-assembly.

A. Cluster equilibration in the lattice gas model

In the self-assembling steady state of the lattice gas model, we count the number of four-bonded particles within each cluster. We average this quantity over clusters of a fixed size n, and we denote this average by $\langle N_4(n)\rangle_{\mathrm{ss}}$. We emphasise that these are the averages over clusters in the self-assembling steady state, and not over the product clusters. It is convenient to compare $\langle N_4(n)\rangle_{\mathrm{ss}}$ with $N_4^{\mathrm{gs}}(n)$, which is the number of four-bonded particles in a cluster of size n that minimises the cluster energy. We then define

$$\langle \Delta n_4(n)\rangle_{\mathrm{ss}} = \frac{1}{n}\left\langle N_4(n) - N_4^{\mathrm{gs}}(n)\right\rangle_{\mathrm{ss}} \qquad (6)$$

to measure the deviation of the cluster "quality" from its ground state value, normalised by the cluster size n.

To test the extent of cluster equilibration, we have performed umbrella sampling, in which we choose a maximal cluster size n_{umb} and reject all MC moves that form clusters of size bigger than n_{umb}. On propagating the dynamics, the system relaxes to a state that satisfies this constraint but is otherwise equilibrated, so that we expect Eq. (5) to hold.[42,59]

In the umbrella-sampled ensemble, we again measure the number of particles with four bonds and average over clusters of size n. The analogue of Eq. (6) within this ensemble is $\langle \Delta n_4(n)\rangle_{\mathrm{umb}} = \frac{1}{n}\langle N_4(n) - N_4^{\mathrm{gs}}(n)\rangle_{\mathrm{umb}}$. Comparison of $\langle \Delta n_4(n)\rangle$ between ensembles allows a test of the cluster equilibration condition: if Eq. (5) holds exactly in the self-assembling steady state then $\langle \Delta n_4(n)\rangle_{\mathrm{ss}} = \langle \Delta n_4(n)\rangle_{\mathrm{umb}}$. In Fig. 5, it can be seen that cluster equilibration holds quite accurately at $\varepsilon_{\mathrm{b}}/T = 2.5$ which is close to the maximum of the yield (recall Figs. 2(b) and 4). However, as $\varepsilon_{\mathrm{b}}/T$ increases and assembly quality is reduced, a strong departure from cluster equilibration is apparent: we find that $\langle \Delta n_4(n)\rangle_{\mathrm{ss}}$ increases while $\langle \Delta n_4(n)\rangle_{\mathrm{umb}}$ decreases. The key point is that the crossover in Q^{prod} in Fig. 4 and the deviations from cluster equilibration occur at similar values of the bond strength. Our conclusion is that effective self-assembly requires transient bond-breaking processes in order to avoid kinetic trapping, and further that these bond-breaking processes need to be frequent enough that the system is close to the cluster equilibration condition (5).

Finally, we note that $\langle \Delta n_4(n)\rangle$ tends to increase with n in a "sawtooth" fashion. The effect is primarily due to the quantity $N_4^{\mathrm{gs}}(n)$ that appears in the definition of $\langle \Delta n_4(n)\rangle$. As n increases, $N_4^{\mathrm{gs}}(n)$ changes in discrete steps of various sizes, depending on the precise nature of the cluster ground state. However, there are often a range of cluster morphologies with energies close to the ground state energy, all of which occur with significant probability in both umbrella-sampled and steady state ensembles. The effect of these clusters is that $\langle N_4(n)\rangle$ depends more smoothly on n than $N_4^{\mathrm{gs}}(n)$, resulting in a sawtooth structure in $\langle \Delta n_4(n)\rangle$. For our purposes, the relevant comparison is between umbrella-sampled and steady-state data, which both exhibit similar n-dependence in this case.

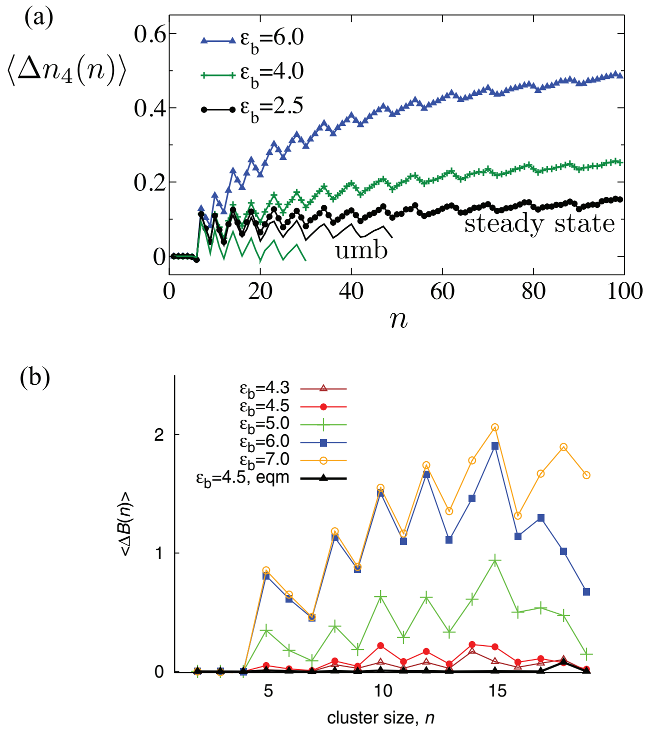

FIG. 5. (a) Measurement of cluster equilibration in the lattice gas model at $\rho = 0.1$, by comparison of steady state and umbrella-sampled ensembles. The data points show $\langle \Delta n_4(n)\rangle_{\mathrm{ss}}$, while the solid lines show $\langle \Delta n_4(n)\rangle_{\mathrm{umb}}$. For $\varepsilon_{\mathrm{b}}/T = 2.5$, the system is far from global equilibrium, but deviations from the cluster equilibration condition are small (compare the black symbols with the solid black lines). As $\varepsilon_{\mathrm{b}}/T$ increases, deviations from cluster equilibration increase. (b) Similar data for the viral capsid model. The deviation in the number of bonds between clusters and their ground states is shown as a function of the cluster size n for indicated values of ε_{b}. All data points correspond to results from the steady state ensemble except for the curve with black ▲ symbols, which were obtained from umbrella sampling.

B. Cluster equilibration in the viral capsid model

To test cluster equilibration in the viral capsid model, we concentrate on the average number of bonds in clusters of size n, denoted by $\langle B(n)\rangle_{\mathrm{ss}}$. We compare this average with the number of bonds $B^{\mathrm{gs}}(n)$ in a cluster of size n with minimal energy. In this case the absolute deviation from the ground state cluster is of particular relevance, since rate equation descriptions of capsid assembly often consider only the ground state morphology for each intermediate size. Therefore, we define

$$\langle \Delta B(n)\rangle_{\mathrm{ss}} = \langle B(n) - B^{\mathrm{gs}}(n)\rangle_{\mathrm{ss}}. \qquad (7)$$

Note that this deviation is not normalised by the cluster size n. As for the lattice gas, we perform umbrella sampling that prohibits the formation of clusters larger than n_{umb}. (Specifically, we use a hybrid Brownian dynamics/Monte Carlo approach where we use a short sequence of unbiased Brownian dynamics steps as a trial move, which is rejected if the size of the largest cluster is greater than n_{umb}.)

Results for $\langle \Delta B(n) \rangle$ are shown in Fig. 5. For the umbrella-sampled data, we find that $\langle \Delta B(n) \rangle_{\mathrm{umb}} \approx 0$ for $\varepsilon_{\mathrm{b}} = 4.5$: this quantity is similarly small for $\varepsilon_{\mathrm{b}} > 4.5$. (The apparent deviation from $\langle \Delta B(n) \rangle_{\mathrm{umb}} \approx 0$ at $n = 18$ in Fig. 5 is likely a result of imperfect equilibration in the umbrella-sampled simulations.) As in the lattice gas data, the steady state measurements show that deviations from cluster equilibration are small near optimal assembly, and grow as kinetic trapping sets in and Q^{prod} decreases.

As in the lattice gas, a sawtooth structure is visible in $\langle \Delta B(n) \rangle_{\mathrm{ss}}$. Here, increasing the cluster size n leads to a change of either one or two bonds in $B^{\mathrm{gs}}(n)$. As $\langle \Delta B(n) \rangle_{\mathrm{ss}}$ deviates from $B^{\mathrm{gs}}(n)$, there are several relevant cluster morphologies in the steady state ensemble which average out the step changes that occur in $B^{\mathrm{gs}}(n)$: typically, the change in $\langle B(n) \rangle_{\mathrm{ss}}$ on increasing n would be somewhere between 1 and 2 bonds. The combination of discrete changes in $B^{\mathrm{gs}}(n)$ and smoother changes in $\langle B(n) \rangle_{\mathrm{ss}}$ results in the apparent sawtooth pattern.

Results in the umbrella-sampled ensemble are analysed in more detail in Appendix C. We find that the free energy of a cluster of size n can be obtained by analysing the total number of bonds formed together with the entropy associated with different ground state morphologies. For the purposes of this section, the sawtooth structure in Fig. 5(b) can be attributed to the fact that some cluster sizes ($n = 5, 8, 10, \ldots$) have ground states in which every capsomer has at least two bonds (see Fig. 6). For these structures $B^{\mathrm{gs}}(n)$ is large, but the number of such morphologies is rather small (see, in particular, Fig. 9(b)). As cluster equilibration breaks down, the effect on $\langle \Delta B(n) \rangle$ is most pronounced for these cluster sizes, since there are many available morphologies with fewer bonds than the ground state, and these morphologies tend to form most quickly as clusters grow. For other cluster sizes, deviations from cluster equilibrium are less pronounced since there are diverse ground state morphologies, all of which are kinetically accessible.

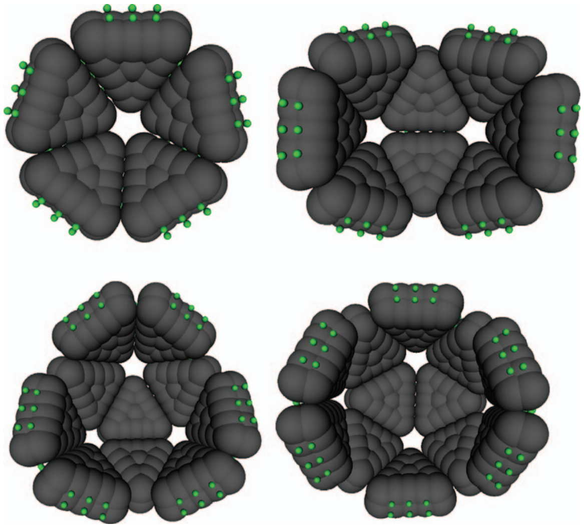

FIG. 6. Clusters of sizes (top left to bottom-right) 5, 8, 10, and 12 in which every capsomer has at least two bonds. For numbers of subunits in between the sizes shown, there are no structures in which every capsomer has at least two unstrained bonds. This pattern gives rise to the sawtooth form for $\langle \Delta B(n) \rangle$ in Fig. 5.

V. KINETIC EQUATIONS IN SELF-ASSEMBLY

In the capsid and lattice gas models, clusters of particles grow as assembly takes place. A natural approach is therefore to describe this process in terms of kinetic rate equations for cluster concentrations. In phase change processes, this idea goes back to Becker and Döring,[25] and a derivation of this approach from the microscopic dynamics of the lattice gas (or Ising) model was considered by Binder and Stauffer.[26] Similar ideas have been developed by Zlotnick and others in order to describe viral capsid assembly.[38,39,43–46]

In this section, we show that non-monotonic steady-state yields Y^{ss} can be predicted by such equations, but we emphasise that these equations fail to capture the decreasing quality Q^{prod} that occurs in both capsid and lattice gas models. We argue that this failure of kinetic rate equations is linked with the breakdown of the cluster equilibration condition (5).

A. Equations for cluster growth and self-assembly

The central idea behind kinetic rate equations is to organise configurations of the system according to the sizes of the clusters that are present in the system. Let $\mathcal{N}_n(t)$ be the number of clusters of size n, at some time t, so that $\rho_n(t) = \mathcal{N}_n(t)/V$ is the concentration of such clusters. For large systems where the various clusters are well-mixed and interact through binary collisions, one often writes

$$\frac{\partial}{\partial t}\rho_n(t) = \sum_{n'}[W^{+}_{n-n',n'}\rho_{n-n'}(t) - W^{+}_{n,n'}\rho_n(t)]\rho_{n'}(t) + \sum_{n'}[W^{-}_{n+n',n'}\rho_{n+n'}(t) - W^{-}_{n,n'}]\rho_n(t), \tag{8}$$

where the coefficients W^{+} and W^{-} are rate constants for binary cluster fusion and cluster fission events, respectively. We use a notation where the sums over n' are unrestricted, but the coefficients $W^{\pm}_{n,n'}$ are zero for $n' \geq n$. For a simple description, we may take $W^{+}_{n,n'}$ and $W^{-}_{n,n'}$ to be finite only when $n' = 1$, recovering the classical Becker-Döring equation.

The restriction to binary collisions may be relaxed straightforwardly (see, for example, Ref. 37) and cases when the clusters are not well-mixed can be treated by field-theoretic approaches.[47] However, an additional assumption on writing Eq. (8) is that *nearly all clusters of size n behave statistically identically, regardless of their shape.* This assumption is tied in with our condition of cluster equilibration above, as discussed in Sec. V C.

B. Non-monotonic production rate *R* in kinetic equations

The steady state ensemble has a natural realisation in terms of these kinetic rate equations. To keep a compact notation, we write $M = n_{\max} - 1$ as the size of the largest clusters that are not removed as products. We consider clusters of sizes $n = 1 \ldots M$, and we restrict ourselves for convenience to monomer binding and unbinding. Then, for $1 < n < M$ we have

$$\frac{\partial}{\partial t}\rho_n(t) = D\rho_1(t)[\rho_{n-1}(t) - \rho_n(t)] + \lambda_{n+1}\rho_{n+1}(t) - \lambda_n\rho_n(t) \tag{9}$$

For simplicity, we have replaced the n-dependent rate constants by a single "diffusion-limited" rate D,[48] and λ_n is the rate for unbinding of a monomer from a cluster of size n. If the system is allowed to equilibrate, we have that $\rho_n^{\mathrm{eq}} = \rho_1^{\mathrm{eq}} \prod_{r=2}^{n}(D\rho_1^{\mathrm{eq}}/\lambda_r)$. In comparing with lattice gas or capsid models, we expect monomer binding and unbinding rates to be related by detailed balance, as

$$D = \lambda_m v \mathrm{e}^{-g_m/T}, \tag{10}$$

where g_m is the (negative) free energy change on monomer binding and v is an entropic factor associated with bonding, with dimensions of volume (specifically, the contribution of monomer attractions to the 2nd virial coefficient of the system is $-v\mathrm{e}^{-g_2/T}$). We expect g_m to be of the order of $-\varepsilon_{\mathrm{b}}$: see Appendix C for explicit calculations for the capsid system.

In the assembling steady state, the equations of motion for $\rho_M(t)$ and $\rho_1(t)$ are modified to include the removal of product clusters: details are given in Appendix B. The total number of particles (subunits) in the system is a constant $\rho_{\mathrm{T}} = \sum_n n\rho_n$. The production rate may also be identified as $R(t) = D\rho_1(t)\rho_M(t)$.

The simplest case is irreversible binding, where bonds never break, so $\lambda_m = 0$ for all m. As shown in Appendix B, the exact result is

$$R^{\infty} = D\rho_1\rho_M = \frac{4D\rho_T^2}{M^2(M+1)^2}. \tag{11}$$

(Within the steady state, we drop all time arguments on ρ_n and R.) The signature of kinetic trapping will be that introducing some non-zero unbinding rates λ_n will lead to an increase in R (holding ρ_T constant). That is, increasing the rate of monomer unbinding increases the production rate R. This is the stalling (starvation) effect of Zlotnick and Endres.[39]

To observe this effect in the steady state, an essential model ingredient is that the unbinding rates λ_m depend on the cluster size m. For simplicity and to maintain contact with Refs. 38 and 39, we suppose that there is a "critical cluster size" m^* above which unbinding is slow $\lambda_m \approx 0$ while for small clusters we take a finite value $\lambda_m = \lambda$. [The critical cluster size should be interpreted in the spirit of classical nucleation theory[26] and the small values of λ_m for large m arise because the binding free energies g_m in Eq. (10) are large and negative.]

The production rate R^{ss} depends on m^*, $n_{\max}$ and a dimensionless parameter $\lambda/D\rho_{\mathrm{T}}$. This last parameter determines the rate of bond-breaking for clusters with $m < m^*$: it is convenient to express this as an "effective bond strength"

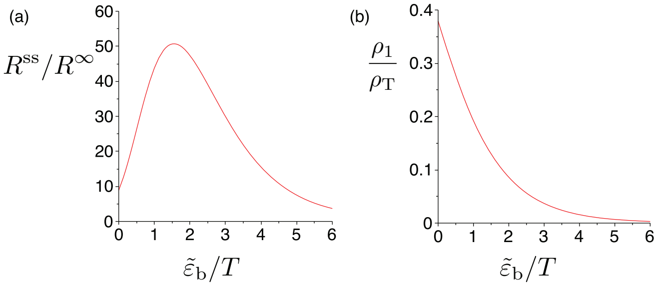

FIG. 7. (a) Rate R^{ss} vs $\tilde{\varepsilon}_{\mathrm{b}}/T$ for kinetic rate equations, showing non-monotonic behaviour due to "kinetic trapping" in states with many intermediates and few monomers. We take $M = 50, m^* = 10$ and the rate is normalised by its value as $\tilde{\varepsilon}_{\mathrm{b}} \to \infty$. (b) The fraction of particles in the assembling steady state that are free monomers, further emphasising that the small rate for large $\tilde{\varepsilon}_{\mathrm{b}}$ arises from the states with a small number of monomers, and hence a small rate of bond formation.

$$\tilde{\varepsilon}_{\mathrm{b}}/T = \log(D\rho_{\mathrm{T}}/\lambda_m). \tag{12}$$

Comparison with Eq. (10) shows that $\tilde{\varepsilon}_{\mathrm{b}} = -g_m - T\log(v\rho_{\mathrm{T}})$ (for $m < m^*$): thus, $\tilde{\varepsilon}_{\mathrm{b}}$ is a grand free energy with ρ_{T} the concentration of a subunit bath. That is, the relevant binding free energy depends on the total subunit density as well as the bonding parameters g_m and v. The key point is that within the rate equation treatment, the full dependence of the system on λ and ρ_{T} can be obtained through the single parameter $\tilde{\varepsilon}_{\mathrm{b}}/T$. (We emphasise, however, that we have assumed that unbinding from large clusters is very slow: the rate λ in this analysis is the rate of unbinding from *small* clusters.)

The central result of this analysis is shown in Fig. 7: the production rate R shows a non-monotonic dependence on $\tilde{\varepsilon}_{\mathrm{b}}$. Since we are working at fixed ρ_{T}, this corresponds to a non-monotonic dependence on ε_{b} in the capsid and lattice gas models. Hence, these results qualitatively mirror the behaviour shown in Fig. 4 as well as the stalling (or starvation) effects discussed by Zlotnick and Endres.[39] We have verified that the non-monotonicity survives on introducing small finite rates for unbinding from large clusters (finite λ_m for $m > m^*$), although a monotonic response is recovered if the unbinding rate is completely independent of m.

Physically, the interpretation of this starvation regime is that a small unbinding rate λ acts to reduce the concentration of monomers ρ_1 since free subunits quickly join growing clusters. The production rate is $R = D\rho_1\rho_M$, so a small concentration ρ_1 reduces this rate strongly. As the unbinding rate λ increases, ρ_1 increases quickly, while the effect on the concentration of large clusters ρ_M is much weaker. Thus, R increases as λ is increased demonstrating that kinetic trapping occurs.

We note that while we have analysed these kinetic equations in the steady state ensemble, similar non-monotonic production rates are observed on starting with disordered states and waiting for clusters to form.[4,38,39,49]

C. Cluster equilibration

The previous results demonstrate that Eq. (8) reproduces one feature of the lattice gas and capsid models, the non-monotonic production rate. However, it is clear from Eq. (8) that this rate equation approach treats all clusters of a given size on the same footing. As discussed above, these approximations are justified if all clusters of a given size behave statistically identically. Classically,[47] the argument supporting this assumption is that large clusters are rare, and that transitions between different morphologies are rapid compared to collisions between clusters. If this separation of time scales holds, one may consider each cluster as a separate subsystem, which relaxes quickly to a "quasiequilibrium" state: the cluster equilibration condition (5) then holds exactly. In practice, the condition of cluster equilibration is much weaker than the assumption of a clear separation of time scales between cluster rearrangement and cluster growth – but the results of Secs. III and IV show that it is the cluster equilibration condition that breaks down as assembly quality falls.

Therefore, when modelling assembly with rate equations of this form, there is an implicit assumption that cluster equilibration holds, and hence that the assembly quality Q^{prod} is independent of temperature. From Figs. 3 and 4, this assumption is not valid once kinetic trapping sets in. Thus, while kinetic rate equations can reproduce a non-monotonic dependence of production rate on bond strength, our results from the steady state ensemble show clearly that these equations miss an important part of the story: the decrease of production quality as bonds get strong.

We note that there are two mechanisms by which cluster equilibration can be violated. In the first, subunits form strong interactions with a sub-optimal number of partners. In other words, each subunit-subunit interaction approximately corresponds to a minimum in the interaction potential, but subunits do not add on to a growing cluster in locations that offer the most interaction partners. In the second mode of violation, subunits form strained bonds which deviate from the ground state of the interaction potential. For example, assembling capsids frequently form hexameric defects, as illustrated in Fig. 3. The first form of cluster equilibration violation can be incorporated into the rate equation approach, at a cost of significantly increased computational complexity, if the space of all possible cluster configurations can be predefined, and then relevant cluster configurations can be enumerated ahead of time[50] or sampled stochastically.[51] However, these approaches have not been used to address the possibility of defective bonds, for which it is not possible to predefine the space of possible cluster configurations.

VI. DISCUSSION AND OUTLOOK

The usefulness of weak interparticle bonds for self-assembly has been commented on by several authors.[4,5,10–14,27,38,52] Thermal fluctuations allow these bonds to be broken: we have shown that this effect can enhance assembly by increasing the concentration of free particles and hence the rate of cluster formation. These results are consistent with the studies by Zlotnick and co-workers. However, our simulations also identify a second mechanism by which weak bonds enhance the assembly of clusters with a given morphology. Namely, bond-breaking processes act to promote cluster equilibration, in the sense of Eq. (5). The qualitative importance of cluster equilibration was first raised by Whitesides and Boncheva;[10] we have attempted to quantify this idea through Eq. (5).

The fact that our findings apply to two very different models suggest that they could apply to a wide variety of systems. However, they need not be completely general. For example care should be taken when interpreting them in the context of one-dimensional assembly such as filament systems (e.g., Ref. 53). Splitting and joining of large oligomers, which can be common in 1D systems,[54] could ameliorate the starvation trap and there are fewer available modes of aberrant assembly in a one-dimensional structure. However, real filaments are not truly one-dimensional and can exhibit polymorphism, branching, or other structural deviations from the ground state (e.g., Refs. 6 and 55).

The importance of kinetic trapping to biological assembly, and the constraints it places on interactions between the constituents, has been vividly demonstrated through experiment (e.g., Refs. 11, 52, and 56) and modelling.[4,5,12,13,15,39,57] If we are to anticipate the design of functionalised particles that self-assemble into ordered structures, the possibility of kinetic trapping must surely be taken into account for these systems as well. In particular, methods for predicting the "optimal weakness" of interparticle bonds could streamline the design process. In Ref. 5, we proposed that the degree of cluster equilibration (or local equilibration) might be measured using fluctuation theorems that couple to the reversibility of bond-formation.

Developments in this direction will be discussed in future articles: here we note that the cluster equilibration condition (5) is weaker than the "local equilibrium" conditions discussed in Ref. 5. For example, Eq. (5) may hold even in the absence of good-mixing conditions, which lead to a deviation from local equilibrium in the sense of Ref. 5. This distinction emphasises the point that, while some degrees of freedom in out-of-equilibrium systems may be locally equilibrated in this sense, other degrees of freedom may be far from equilibrium. For example, the recent results of Russo and Sciortino[58] seem to indicate that density fluctuations are much closer to a local equilibrium distribution than energy fluctuations. We conjecture that the near-local equilibration of density is linked with a weak violation of the good-mixing assumption, while the energy fluctuations reflect a stronger violation of cluster equilibration, in this out-of-equilibrium system.

More generally, we conclude that our results are entirely consistent with the general idea[10] that effective self-assembly occurs through the reversible formation of numerous weak bonds. We have used statistical mechanical methods including the steady state ensemble and comparison with umbrella-sampled data in order to test this idea quantitatively, with a view to exploiting it in the design and control of self-assembly process. In particular, the breakdown of cluster equilibration when bonds are strong is a kinetic effect that is not taken into account in classical theories of self-assembly and phase change. We believe that the development of other quantitative methods for characterising this effect is a key challenge for theoretical studies of self-assembly, and we look forward to further progress in this area.

ACKNOWLEDGMENTS

We thank Steve Whitelam, Phillip Geissler, and David Chandler for many discussions on the importance of reversibility in self-assembly, and R.L.J. thanks Stephen Williams for helpful discussions of local equilibration and quasi-equilibrium. This work was supported by Award No. R01AI080791 from the National Institute Of Allergy And Infectious Diseases (to MFH and OME) and by the Engineering and Physical Sciences Research Council (UK) through Grant Nos. EP/G038074/1 and EP/I003797/1 (to R.L.J.). M.F.H. also acknowledges support by the National Science Foundation (NSF) through the Brandeis Materials Research Science and Engineering Center (MRSEC). Computational resources were provided by the NSF through TeraGrid computing resources (specifically the Purdue Condor pool) and the Brandeis HPCC.

APPENDIX A: DESCRIPTION OF THE CAPSID MODEL

The model subunits are comprised of a set of overlapping spherical "excluders" that enforce excluded volume and spherical "attractors" with short-range pairwise, complementary attractions that decorate the binding interfaces of the subunit. Each subunit has two layers of excluders and attractors. Attractor positions are arranged so that complementary attractors along a subunit-subunit interface perfectly overlap in the ground state configuration; excluders on either side of the interface are separated by exactly the cut off of their potential (x_c, Eq. (A2)). Subunits have no internal degrees of freedom—they translate and rotate as rigid bodies.

The capsid subunits interact through a pairwise potential, which can be decomposed into pairwise interactions between the elemental building blocks—the excluders and attractors. The potential of capsomer subunit i, $U_{\mathrm{cap},i}$, with position $\boldsymbol{R}_i$, attractor positions $\{\boldsymbol{a}_i\}$ and excluder positions $\{\boldsymbol{b}_i\}$ interacting with subunit j is the sum of a repulsive potential between every pair of excluders and an attractive interaction between complementary attractors:

$$\begin{aligned}
&U_{\mathrm{cc}}(\boldsymbol{R}_i, \{\boldsymbol{a}_i\}, \{\boldsymbol{b}_i\}, \boldsymbol{R}_j, \{\boldsymbol{b}_j\}, \{\boldsymbol{a}_j\}) \\
&\quad = \sum_{k,l}^{N_\mathrm{b}} \mathcal{L}_8\left(\left|\boldsymbol{R}_i + \boldsymbol{b}_i^k - \boldsymbol{R}_j - \boldsymbol{b}_j^l\right|,\ 2^{1/4}\sigma_\mathrm{b},\ \sigma_\mathrm{b}\right) \\
&\qquad + \sum_{k,l}^{N_\mathrm{a}} \chi_{kl}\varepsilon_\mathrm{b}\mathcal{L}_4\left(\left|\boldsymbol{R}_i + \boldsymbol{a}_i^k - \boldsymbol{R}_j - \boldsymbol{a}_j^l\right| - 2^{1/2}\sigma_\mathrm{a}, 4\sigma_\mathrm{a}, \sigma_\mathrm{a}\right),
\end{aligned} \tag{A1}$$

where ε_b is an adjustable parameter setting the strength of the capsomer-capsomer attraction at each attractor site, N_b and N_a are the number of excluders and attractors, respectively, σ_b and σ_a are the diameters of the excluders and attractors, $\boldsymbol{b}_i^k$ ($\boldsymbol{a}_i^k$) is the body-centered location of the k^{th} excluder (attractor) on the ith subunit, χ_{kl} is 1 if attractors k and l are overlapping in a completed capsid (Figure 8) and 0 otherwise. The diameter of attractors is set to $\sigma_\mathrm{a} = \sigma_\mathrm{b}/5$ for all results in this work. The function $\mathcal{L}_p$ is defined as a truncated Lennard-Jones-like potential:

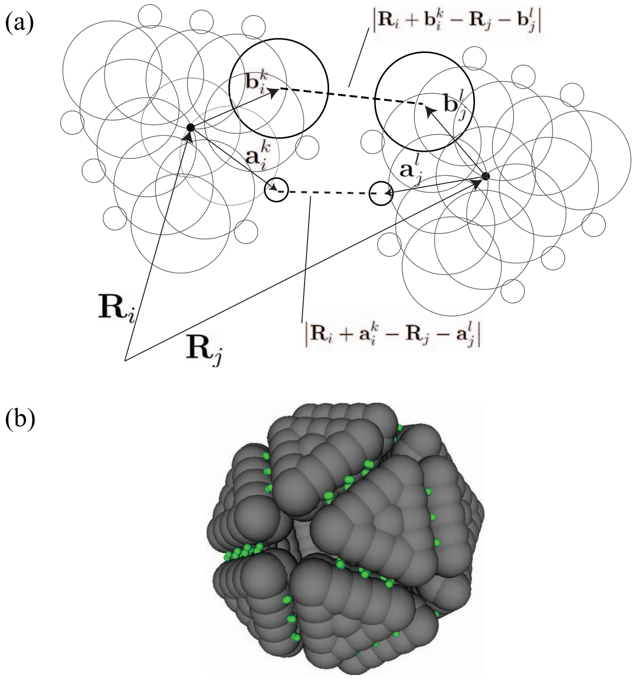

FIG. 8. The model capsid geometry. (a) Two dimensional projection of one layer of a model subunit illustrating the geometry of the capsomer-capsomer pair potential, Eq. (A1), with a particular excluder and attractor highlighted from each subunit. The potential is the sum over all excluder-excluder and complementary attractor-attractor pairs. (b) An example of a well-formed model capsid from a simulation trajectory.

$$\mathcal{L}_p(x, x_\mathrm{c}, \sigma) \equiv \begin{cases} 4\left(\left(\frac{x}{\sigma}\right)^{-p} - \left(\frac{x}{\sigma}\right)^{-p/2}\right) + 1; & x < x_\mathrm{c} \\ 0; & \text{otherwise}, \end{cases} \tag{A2}$$

In our dynamical simulations, the capsomer subunits have anisotropic translational and rotational diffusion constants, calculated as in Ref. 30. The unit of time is set by the diffusion constant of a single excluder D, and we define $t_0 = \sigma_\mathrm{b}^2/D$. In these dimensionless units, the eigenvalues of the translational and rotational diffusion tensors for capsomer subunits are $\{0.283, 0.283, 0.197\}$ and $\{0.1906, 0.1906, 0.0984\}$, respectively.

APPENDIX B: PRODUCTION RATE WITHIN THE STEADY STATE

Here, we explain how we solved the kinetic equations (9) to obtain the cluster production rate R in the steady state ensemble. As discussed in the main text, Eq. (9) with $n = M$ reduces to

$$\frac{\partial}{\partial t}\rho_M(t) = D\rho_1(t)[\rho_{M-1}(t) - \rho_M(t)] - \lambda_M \rho_M(t) \tag{B1}$$

and the production rate is

$$R(t) = D\rho_1(t)\rho_M(t). \tag{B2}$$

For completeness, we also give the equation of motion for the monomer concentration $\rho_1(t)$ within the steady state, which is

$$\frac{\partial}{\partial t}\rho_1(t) = MR(t) - 2D\rho_1(t)^2 + 2\lambda_2\rho_2(t) + \sum_{n=2}^{M}[\lambda_n\rho_n(t) - D\rho_1(t)\rho_{n-1}(t)]. \qquad \text{(B3)}$$

In the following, we work in the steady state so we suppress all time dependence of the ρ_n. Equation (B2) gives $\rho_M = D\rho_1/R$ while Eq. (B1) gives $D\rho_{M-1} = (R/\rho_1)(1 + (\lambda_M/D\rho_1))$. The remaining ρ_n may then be obtained inductively since Eq. (9) reduces to $D(\rho_n - \rho_{n-1}) = (1/\rho_1)[\lambda_{n+1}\rho_{n+1} - \lambda_n\rho_n]$, so that ρ_{n-1} is given in terms of ρ_m with $m \geq n$. For $1 \leq n \leq M-2$ we arrive at

$$D\rho_n = \frac{R}{\rho_1} + \frac{1}{\rho_1}\sum_{m=n+1}^{M-1}[\lambda_m\rho_m - \lambda_{m+1}\rho_{m+1}], \qquad \text{(B4)}$$

which allows calculation of all of the ρ_n, in terms of R, ρ_1, and λ_n.

A simple case is when no unbinding takes place, so that $\lambda_m = 0$. Then, $\rho_n = \rho_1$ for all n, and $\rho_T = M(M+1)\rho_1/2$. Hence, the production rate for irreversible binding is given by Eq. (11).

We now turn to the problem described in the main text, where $\lambda_m = \lambda$ for $m \leq m^*$, with $\lambda_m = 0$ for $m > m^*$. It then follows from (B4) that,

$$\rho_n = \begin{cases} \frac{R}{D\rho_1}, & n \geq m^*, \\ \frac{R}{D\rho_1} S(\tilde{\lambda}, m^* - n), & n < m^*, \end{cases} \qquad \text{(B5)}$$

where $S(x,n) = (1 - x^{n+1})/(1-x)$ is obtained by summing a geometrical progression and $\tilde{\lambda} = \lambda/D\rho_1$. We then sum over n to obtain ρ_{T} and eliminate R from the result using

$$R = D\rho_1^2/S(\tilde{\lambda}, m^* - 1), \qquad \text{(B6)}$$

which follows from Eq. (B5). The result is

$$(D\rho_T/\lambda) = \frac{(M - m^*)(M + m^* + 1) + f(\tilde{\lambda}, m^*)}{2\tilde{\lambda}S(\tilde{\lambda}, m^* - 1)}, \qquad \text{(B7)}$$

with

$$f(\tilde{\lambda}, m) = \left[m(m+1) - 2m\frac{\partial}{\partial\tilde{\lambda}} + \frac{\partial^2}{\partial\tilde{\lambda}^2}\right] S(\tilde{\lambda}, m). \qquad \text{(B8)}$$

[We used $\sum_{r=1}^{m} rx^r = x(\partial/(\partial x))S(x,m)$ and similarly $\sum_{r=2}^{m} r(r-1)x^r = x^2(\partial^2/\partial x^2)S(x,m)$.] Dimensional analysis shows that the normalised rate R/R^∞ depends only on M, m^*, and $\lambda/D\rho_{\mathrm{T}}$. We, therefore, fix these parameters and solve Eq. (B7) numerically for $\tilde{\lambda}$, obtaining the monomer concentration $\rho_1 = \lambda/(D\tilde{\lambda})$. The rate R may then be calculated from Eq. (B6), as shown in Fig. 7 and discussed in the main text.

APPENDIX C: BINDING FREE ENERGIES FOR THE CAPSID MODEL

We define the free energy change on adding a capsomer to a cluster of size $n-1$ (that is, the capsomer binding free energy) to be

$$g_{\mathrm{b}}(n) = -T\ln\left[\frac{\rho_n}{\rho_{n-1}}\frac{c_{\mathrm{ss}}}{\rho_1}\right], \qquad \text{(C1)}$$

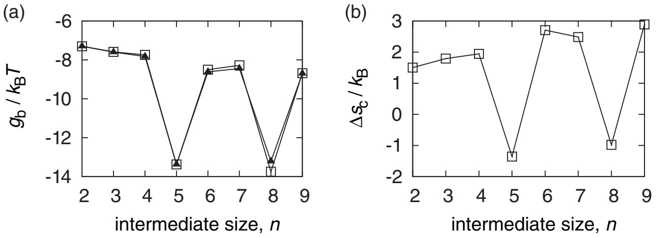

FIG. 9. (a) The binding free energy g_{b} to add an additional subunit is shown as a function of intermediate size for $\varepsilon_{\mathrm{b}} = 4.5$. The ▲ symbols denote values computed from umbrella sampling simulations, while the □ symbols were calculated based on the cluster configurational entropy, as described in the text. (b) The change in the configurational entropy, Δs_{c}, computed from the ground state cluster geometries is shown as a function of intermediate size.

where ρ_n is the concentration of clusters of size n (see Sec. V A) and c_{ss} is a reference concentration (always required when quoting binding free energies). We take $k_{\mathrm{B}} = 1$ throughout this section. Following Ref. 27, by comparing the size of our capsid to the size of a satellite tobacco mosaic virus capsid, we assign $c_{\mathrm{ss}} = 8\sigma_{\mathrm{b}}^{-3}$ to correspond to 1 M. In Fig. 9(a), we show binding free energies at $\varepsilon_{\mathrm{b}}/T = 4.5$, obtained from umbrella-sampled computer simulations as described in the main text.

All of our results for the capsid model are in the dilute regime where the system may be considered as a non-interacting gas of clusters, so the $g_{\mathrm{b}}(n)$ are independent of the overall capsomer density. However, these free energies do depend on the bond strength ε_{b}. We find that the free energy of dimerisation is approximately linear in ε_{b} over the range we consider,

$$g_{\mathrm{b}}(2) \approx -3.5\varepsilon_{\mathrm{b}} - Ts_{\mathrm{b}} - T\Delta s_{\mathrm{c}}(2), \qquad \text{(C2)}$$

with the binding entropy penalty, $s_{\mathrm{b}} = -9.95$ and the configurational entropy change for dimerisation, $\Delta s_{\mathrm{c}}(2) = \ln(9/2) \approx 1.5$. Here, $\Delta s_{\mathrm{c}}(2)$ is an example of a difference in "configurational entropy": we define $\Delta s_{\mathrm{c}}(n) = \ln(\Omega_n/\Omega_{n-1})$, with Ω_n is the number of distinct ground state cluster configurations with n subunits. (In counting distinct configurations, the three edges of each capsomer are assumed to be distinguishable, but configurations related by global rotations are not distinct from one another. So the number of distinct dimer configurations is $\Omega_2 = (3^2/2)$ since there are three possible binding sites on each capsomer (hence 3^2 configurations) while the factor of 2 accounts for a rotation symmetry of the entire dimer.) Note that the value for s_{b} given in Ref. 27 contains a typo.

The binding free energy depends on the number of contacts that can be formed and the symmetry of the ground state complex—we now analyse these effects by explicit counting arguments. We calculate $\Delta s_{\mathrm{c}}(n)$ geometrically for n between 2 and 9: the approach follows Zlotnick[43] except that we consider all possible ground state structures. Results are shown in Fig. 9(b). Then, to calculate the "theoretical" binding energies (□ symbols) in Fig. 9(a), we write $g_{\mathrm{b}}(n) = g_{\mathrm{b}}^{\mathrm{int}}(n) - T\Delta s_{\mathrm{c}}(n)$, where $g_{\mathrm{b}}^{\mathrm{int}}$ is the "interaction part" of the capsomer binding free energy (see Eq. (C2)). To approximate this quantity, we assume that if the capsomer that binds to the cluster gains just one bond then $g_{\mathrm{b}}^{\mathrm{int}}(n) \approx g_{\mathrm{b}}^{\mathrm{int}}(2)$, while if the capsomer gains two bonds then

$g_b^{\text{int}}(n) \approx g_b^{\text{int}}(5)$. Thus, taking the values of $g_b^{\text{int}}(2)/T = -5.8$ and $g_b^{\text{int}}(5)/T = -14.7$ obtained from umbrella sampling at $\varepsilon_b/T = 4.5$ together with the calculated values of Ω_n, we are able to obtain $g_b(n)$ for $n \leq 9$, as shown in Fig. 9(a). The fit to the directly measured free energies is excellent.

To obtain the critical capsomer concentration shown in Fig. 1, we extrapolate this procedure to a full capsid. We define the capsid free energy,

$$G_b(20) = -T \ln \left[\frac{\rho_{20} c_{ss}^{19}}{\rho_1^{20}} \right]. \tag{C3}$$

Once this quantity is known, and under the (excellent) assumption that the equilibrium state of the system is dominated by a combination of monomers and complete capsids with very few intermediate-sized clusters, the critical capsomer concentration is obtained by setting $\rho_1 = 20\rho_{20}$, so that half of all capsomers are in complete capsids and taking the total capsomer density as $\rho_{cc} = 2\rho_1$. Hence,

$$\rho_{cc} = \frac{2c_{ss}}{20^{1/19}} \exp\left(\frac{G_b(20)}{19T} \right). \tag{C4}$$

From the definitions of $g_b(n)$ and Ω_n, the capsid free energy can be written as

$$G_b(20) = \sum_{n=2}^{20} g_b(n) = -T \ln \Omega_{20} + \sum_{n=2}^{20} g_b^{\text{int}}(n), \tag{C5}$$

where the combinatorial factor associated with the icosahedral capsid is $\Omega_{20} = 3^{20}/60$. We assume that the dependence of g_b on ε_b is that $g_b^{\text{int}}(2)/T \approx -3.5\varepsilon_b/T + s_b$ as discussed above while $g_b^{\text{int}}(5)/T \approx -7\varepsilon_b/T + s_b'$ since the capsomer that binds makes two bonds in this case. We also require an approximation for $g_b^{\text{int}}(20)$: adding the final capsomer involves adding three new bonds so we approximate this as $g_b^{\text{int}}(20) \approx g_b^{\text{int}}(5) - 3.5\varepsilon_b$, including the extra energy gained from the third bond and neglecting any extra entropy lost. Hence,

$$G_b(20) \approx -T \log \Omega_{20} + 9g_b^{\text{int}}(2) + 10g_b^{\text{int}}(5) - 3.5\varepsilon_b, \tag{C6}$$

and using the numerical values for $g_b^{\text{int}}(2)$ and $g_b^{\text{int}}(5)$ obtained from umbrella sampling at $\varepsilon_b/T = 4.5$, we arrive at $G_b(20)/T = -233 - 105(\varepsilon_b/T - 4.5)$. Finally, using (C4) and (C6) together gives the result for the critical capsomer concentration shown in Fig. 1(a).

[1] D. L. D. Caspar and A. Klug, Cold Spring Harbor Symp. Quant. Biol. **27**, 1 (1962).

[2] H. Fraenkel-Conrat and R. C. Williams, Proc. Natl. Acad. Sci. U.S.A. **41**, 690 (1955); A. Klug, Philos. Trans. R. Soc. London, Ser. B **354**, 531 (1999); A. Zlotnick, R. Aldrich, J. M. Johnson, P. Ceres, and M. J. Young, Virology **277**, 450 (2000); J. Sun, C. DuFort, M. Daniel, A. Murali, C. Chen, K. Gopinath, B. Stein, M. De, V. M. Rotello, A. Holzenburg, C. Kao, and B. Dragnea, Proc. Natl. Acad. Sci. U.S.A. **104**, 1354 (2007).

[3] A. Zlotnick, J. Mol. Biol. **241**, 59 (1994); B. Berger, P. W. Shor, L. Tucker-Kellogg, and J. King, Proc. Natl. Acad. Sci. U.S.A. **91**, 7732 (1994); T. Chen, Z. Zhang, and S. C. Glotzer, Proc. Natl. Acad. Sci. U.S.A. **194**, 717 (2007); H. D. Nguyen, V. S. Reddy, and C. L. Brooks III, Nano Lett. **7**, 338 (2007).

[4] M. F. Hagan and D. Chandler, Biophys. J. **91**, 42 (2006).

[5] R. L. Jack, M. F. Hagan, and D. Chandler, Phys. Rev. E **76**, 021119 (2007).

[6] Y. Yang, R. Meyer, and M. F. Hagan, Phys. Rev. Lett. **104**, 258102 (2010).

[7] S. Whitelam, Phys. Rev. Lett. **105**, 088102 (2010).

[8] P. Rothemund, Nature (London) **440**, 297 (2006).

[9] S. Sacanna, W. T. M. Irvine, P. M. Chaikin, and D. J. Pine, Nature (London) **464**, 575 (2010).

[10] G. M. Whitesides and M. Boncheva, Proc. Natl. Acad. Sci. U.S.A. **99**, 4769 (2002).

[11] A. Zlotnick, J. Mol. Biol. **366**, 14 (2007).

[12] D. C. Rapaport, Phys. Rev. Lett. **101**, 186101 (2008).

[13] O. M. Elrad and M. F. Hagan, Nano Lett. **8**, 3850 (2008).

[14] S. Whitelam, E. H. Feng, M. F. Hagan, and P. L. Geissler, Soft Matter **5**, 1251 (2009).

[15] H. D. Nguyen, V. S. Reddy, and C. L. Brooks III, Nano Lett. **7**, 338 (2007).

[16] A. W. Wilber, J. P. K. Doye, A. A. Louis, E. G. Noya, M. A. Miller, and P. Wong, J. Chem. Phys. **127**, 085106 (2007).

[17] A. W. Wilber, J. P. K. Doye, and A. A. Louis, J. Chem. Phys. **131**, 175101 (2009).

[18] D. Klotsa and R. L. Jack, Soft Matter **6**, 6294 (2011).

[19] C. Allain, M. Cloitre, and M. Wafra, Phys. Rev. Lett. **74**, 1478 (1995).

[20] P. J. Lu, E. Zaccarelli, F. Ciulla, A. B. Schofield, F. Sciortino, and D. A. Weitz, Nature (London) **453**, 499 (2008).

[21] D. Wales, *Energy Landscapes* (Cambridge University Press, Cambridge, England, 2004).

[22] F. Baletto, J. P. K. Doye, and R. Ferrando, Phys. Rev. Lett. **88**, 075503 (2002).

[23] C. P. Royall, S. R. Williams, T. Ohtsuka, and H. Tanaka, Nature Mater. **7**, 556 (2008).

[24] P. Meakin, Phys. Rev. Lett. **51**, 1119 (1983).

[25] R. Becker and W. Döring, Ann. Phys. (Leipzig) **416**, 719 (1935).

[26] K. Binder and D. Stauffer, Adv. Phys. **25**, 343 (1976).

[27] O. Elrad and M. F. Hagan, Phys. Biol. **7**, 045003 (2010).

[28] D. Rapaport, J. Johnson, and J. Skolnick, Comput. Phys. Commun. **121–122**, 231 (1999); D. Rapaport, Phys. Rev. E **70**, 051905 (2004).

[29] A. Branka and D. Heyes, Phys. Rev. E **60**, 2381 (1999); D. Heyes and A. Branka, Mol. Phys. **98**, 1949 (2000).

[30] J. G. de la Torre and B. Carrasco, Biopolymers **63**, 163 (2002).

[31] The simulations with $\varepsilon_b = 4.1$ and $\varepsilon_b = 4.3$ have not completely transitioned into the logarithmic growth phase at this time point. Since the nucleation time rises exponentially with decreasing ε_b (Ref. 38); the transition to logarithmic growth phase increases in a similar fashion. However, the variation of yield with respect to ε_b is robust to this choice of observation time.

[32] The maximal interaction energy between subunits is $6\varepsilon_b$, so the value of ε_b at optimal assembly may seem large. However, there is a significant entropy penalty on binding, due to the short length scale of the interactions and their orientation-dependence. Thus with $\varepsilon_b = 4.5$, the binding free energy is approximately $g_b = -7k_BT$, as discussed in Appendix C and Fig. 9.

[33] S. Whitelam and P. L. Geissler, J. Chem. Phys. **127**, 154101 (2007).

[34] A. Bhattacharyay and A. Troisi, Chem. Phys. Lett. **458**, 210 (2008).

[35] If we had taken Kawasaki dynamics where MC attempted moves involve single particles, then the diffusion constant of clusters of particles would be very small when ε_b is large. The resulting dynamics would not be consistent with the physical scenario considered here.

[36] R. J. Baxter, *Exactly Solved Models in Statistical Mechanics* (Academic, New York, 2002).

[37] L. Maibaum, Phys. Rev. Lett. **101**, 256102 (2008).

[38] M. F. Hagan and O. M. Elrad, Biophys. J. **98**, 1065 (2010).

[39] D. Endres and A. Zlotnick, Biophys. J. **83**, 1217 (2002).

[40] We note that Y_4^{ss} depends on the product size cutoff $n_{\max}$ while $n_4(t)$ depends on the time t, so the comparison between these measurements is necessarily qualitative, but their respective dependencies on $n_{\max}$ and t are weak and they show similar non-monotonic behaviour.

[41] In general, such rates depend on the diffusion constants and sizes of the relevant clusters but including such factors does not change any qualitative features of our results.

[42] In Ref. 5, some of us argued that the poor assembly that often occurs in system with strong intercomponent bonds is related to the breakdown of a condition that we called local equilibration (our use of this term is similar in spirit to an analogous condition in non-equilibrium thermodynamics (Ref. 59), but the locality that we refer to is the space of cluster configurations, rather than in the spatial coordinates of the system). Here, we use the term "cluster equilibrium" to formulate a closely related condition and we test the extent to which this condition is correlated with effective self-assembly.

[43] A. Zlotnick, J. Mol. Biol. **241**, 59 (1994).

[44] A. Zlotnick, J. Mol. Recognit. **18**, 479 (2005).

[45] A. Y. Morozov, R. F. Bruinsma, and J. Rudnick, J. Chem. Phys. **131** (2009).

[46] M. F. Hagan, J. Chem. Phys. **130**, 114902 (2009).

[47] A. Bray, Adv. Phys. **43**, 357 (1994).

[48] In principle, measurements of $\mathcal{N}_{n,\alpha}(t)/\mathcal{N}_{n,\gamma}(t)$ in the umbrella-sampled ensemble may depend on the value of n_{umb}, especially if morphologies α and γ have different excluded volumes. In this case the cluster equilibration condition would be ill-defined due to cluster-cluster interactions. However, we used a range of values for n_{umb} in our simulations and we did not observe any such dependence.

[49] A. Zlotnick, J. M. Johnson, P. W. Wingfield, S. J. Stahl, and D. Endres, Biochemistry **38**, 14644 (1999).

[50] P. Moisant, H. Neeman, and A. Zlotnick, Biophys. J. **99**, 1350 (2010).

[51] B. Sweeney, T. Zhang, and R. Schwartz, Biophys. J. **94**, 772 (2008).

[52] P. Ceres and A. Zlotnick, Biochemistry **41**, 11525 (2002).

[53] S. Whitelam, C. Rogers, A. Pasqua, C. Paavola, J. Trent, and P. L. Geissler, Nano Lett. **9**, 292 (2009).

[54] T. P. J. Knowles, C. A. Waudby, G. L. Devlin, S. I. A. Cohen, A. Aguzzi, M. Vendruscolo, E. M. Terentjev, M. E. Welland, and C. M. Dobson, Science **326**, 1533 (2009).

[55] M. Fandrich, J. Meinhardt, and N. Grigorieff, Prion **3**, 89 (2009).

[56] S. P. Katen, S. R. Chirapu, M. G. Finn, and A. Zlotnick, ACS Chem. Biol. **5**, 1125 (2010).

[57] S. D. Hicks and C. L. Henley, Phys. Rev. E **74**, 031912 (2006).

[58] J. Russo and F. Sciortino, Phys. Rev. Lett. **104**, 195701 (2010).

[59] S. de Groot and P. Mazur, *Non-equilibrium Thermodynamics* (Dover, Mineola, New York, 1984).