Fluctuation-dissipation ratios in the dynamics of self-assembly

Robert L. Jack,[1] Michael F. Hagan,[1,2] and David Chandler[1]
[1]*Department of Chemistry, University of California at Berkeley, Berkeley, California 94709, USA*
[2]*Department of Physics, Brandeis University, Waltham, Massachusetts 02454, USA*
(Received 24 May 2007; published 22 August 2007)

We consider two seemingly very different self-assembly processes: formation of viral capsids and crystallization of sticky disks. At low temperatures, assembly is ineffective, since there are many metastable disordered states, which are a source of kinetic frustration. We use fluctuation-dissipation ratios to extract information about the degree of this frustration. We show that our analysis is a useful indicator of the long-term fate of the system, based on the early stages of assembly.

DOI: 10.1103/PhysRevE.76.021119 PACS number(s): 05.40.−a, 81.16.Dn, 87.10.+e

I. INTRODUCTION

Self-assembly processes can be loosely defined as those in which simple building blocks assemble spontaneously into highly ordered structures. Assembly is of vital importance in biology, where cells use dynamically assembled protein structures to control the shapes of lipid membranes [1–3], and the life cycle of viruses involves spontaneous assembly of their protein coats [4,5]. Recently, self-assembly has also been used to develop nanostructured materials [6,7], which often draw inspiration from biological systems. The self-assembly of viral capsids [8] has been the subject of elegant experimental [4] and theoretical studies [5,9,10]. In this paper, we study a model [9] designed to mimic this assembly process. At low temperatures, assembly is frustrated by the presence of long-lived disordered states. The avoidance of this frustration is crucial for successful assembly. This effect is rather general, as we illustrate by also considering the formation of ordered structures in a two-dimensional system of sticky disks. We analyze the crossover between frustrated and unfrustrated regimes, and show that fluctuation-dissipation ratios (FDRs) [11–14] associated with the early stages of assembly are correlated with the long-time yield of these processes. This represents an additional application of FDRs, which have been discussed previously in the context of glassy dynamics. We discuss how this analysis might be helpful in the design of self-assembling systems.

In general, successful self-assembly requires both that the equilibrium state of the system is an ordered structure, and that the system reaches this ordered state in the time available for the biological or experimental application. The first condition is thermodynamic: for the systems studied here, the low-energy ordered states are known, and this crossover can be estimated by free energy arguments, as in [9]. (Note, however, that if "liquidlike" states are relevant near the thermodynamic crossover, then this will lead to more complicated behavior, as in [10].) The second condition for successful assembly is kinetic in origin: it is illustrated for a model system of viral capsid assembly in Fig. 1. The degree of assembly shows a maximum at a finite temperature T^*. As the temperature is lowered through T^*, the ordered state becomes more probable at equilibrium, but the self-assembly process becomes less and less effective: we refer to this change as a "kinetic crossover."

The purpose of this paper is to use dynamical observables to study the behavior near T^*. Since this is the regime of most efficient assembly, it is relevant both biologically and for applications of self-assembly in nanoscience. While the kinetic crossover can always be identified by measuring the degree of assembly, as in Fig. 1, achieving this in a computer simulation requires access to long time scales, which restricts the range of systems that can be studied. In this paper, we show how FDRs can be used to locate the kinetic crossover using simulations on relatively short time scales. (It is necessary to average over many such short simulations, but such averaging is trivially parallelizable.) We also discuss how these response functions might be measured experimentally in ordering processes that occur on complex energy landscapes.

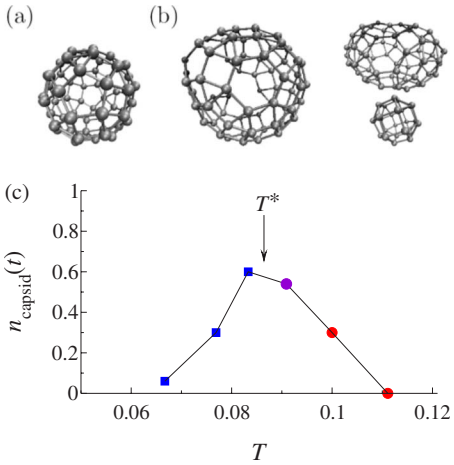

FIG. 1. (Color online) Assembly of model capsids in the B_4 model of Ref. [9]. (a) A well-formed model capsid, with icosahedral symmetry. (b) Representative selection of metastable states formed at reduced temperature $T=0.067$ and reduced time $t=3\times10^5$ (see the text for definitions). (c) Plot of the capsid yield at $t=3\times10^5$, which is nonmonotonic in the reduced temperature. The yield is the fraction of particles in complete capsids, identified as in [9]. Here and throughout, red and blue symbols identify high and low temperatures, respectively. We also indicate the approximate location of the kinetic crossover, at reduced temperature T^*.

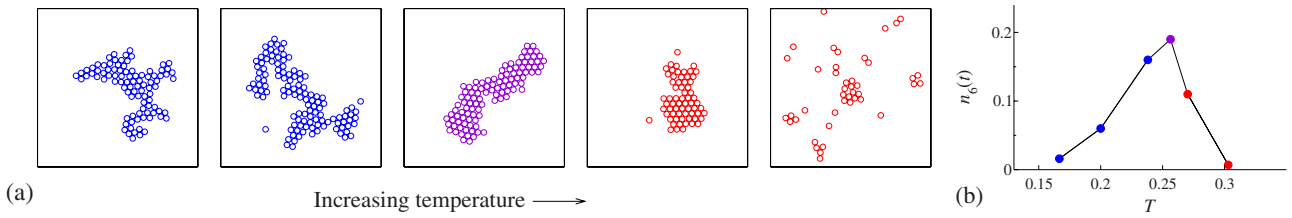

FIG. 2. (Color online) Assembly of sticky disks. (a) Typical part-assembled structures at reduced time $t=5\times10^6$, and reduced temperatures $T=0.17,0.2,0.26,0.27,0.33$, from left to right. Illustrated regions are of size $25a_0\times25a_0$. The crystallinity is poor at low temperatures, due to the metastability of the disordered states. (b) Plot of the fraction of particles with six bonds, which is a measure of the yield of the assembly process. Compare Fig. 1(c).

II. MODELS

A. Model capsids

The first model that we discuss describes the assembly of viral capsids. Full details are given in [9]. The model consists of rigid subunits, the "capsomers," which interact by isotropic repulsive forces, and directional attractions. The low-energy states in the model contain "capsids," each of which consists of 60 subunits in a cage structure, with icosahedral symmetry. We use the B_4 variant of this model, which means that the attractive potential favors the capsid structure shown in Fig. 1(a). The subunit diameter is σ, and the density of subunits is ρ. The parameters of the model are the reduced capsomer density $\rho\sigma^3$ and the reduced temperature T (measured in units of ε_b/k_B, where ε_b is the energy associated with the attractive potential and k_B is Boltzmann's constant). In addition, the specificity of the directional attractions is controlled by the angular parameters θ_m and ϕ_m. The data of this paper are obtained under the representative conditions $\rho\sigma^3=0.11$, $\theta_m=1.5$, and $\phi_m=3.14$. We simulate a system of 1000 capsomers in a cubic box with periodic boundaries. The capsomers evolve according to overdamped Brownian dynamics, and the unit of time is $(\sigma^2/48D)$, where D is the capsomer diffusion constant. The rotational diffusion constant of each capsomer is $D_r=2.5(D/\sigma^2)$, as in [9].

B. Sticky disks

We also consider a second model whose subunits are sticky disks which interact by an attractive square-well potential of depth J and range a_0, and a repulsive hard core of range $0.9a_0$. We quench the system into the solid-vapor phase coexistence regime, so that the equilibrium state has most of the disks in a single close-packed crystallite. However, we use Monte Carlo dynamics that are chosen to accentuate the effects of kinetic frustration. We move bonded clusters as rigid bodies, allowing translation and rotation, but no internal rearrangements. To reflect the slow motion of large clusters, we use an average translational step size of $0.1(a_0/M)$ and a rotational step of $\pi/(10I)$ rad, where M is the number of particles in the cluster and Ia_0^2 its moment of inertia (in units of the disk mass). The reduced time t is measured in Monte Carlo sweeps, and the reduced temperature T is measured in units of J/k_B. Clusters can rearrange only by bond breaking. These events are sampled by the cleaving algorithm of [15], with equal fictitious and physical temperatures. It is an off-lattice generalization of the Swendsen-Wang algorithm [16]. At low temperatures, the dynamics mean that bonds are broken very rarely, and aggregation of the disks is diffusion limited. At $T=0$, the system reduces to diffusion-limited cluster aggregation [17].

The crossover from effective to ineffective assembly in the capsid system was shown in Fig. 1. We show similar results for the disk system in Fig. 2. The system contains 400 disks in a periodic square box of side $100a_0$. The system does not reach full phase separation into close-packed structures on the time scales accessible to our simulation, so all of our data are in the out-of-equilibrium regime.

III. FLUCTUATION-DISSIPATION RATIOS

The nonmonotonic yields shown in Figs. 1(c) and 2(b) mean that for the observation times considered, and when the temperature is small, reducing the temperature does not result in a decrease in of the total energy. This kind of negative response to temperature perturbations is familiar in systems with activated dynamics [14]. In the self-assembling systems considered here, the kinetic crossover at T^* is associated with a change from positive to negative response on the long time scales considered in Figs. 1(c) and 2(b). In this section, we use fluctuation-dissipation ratios to show that the crossover between positive and negative response has signatures that can be detected on much shorter time scales.

A. Basic idea

Fluctuation-dissipation ratios (sometimes also called correlation-response ratios) have been widely studied in the context of aging of glassy systems [11]. Imagine applying an instantaneous perturbation to a single subunit (disk or capsomer) at a time t_w, and measuring the effect of this perturbation at some later time t. For a system at equilibrium, the fluctuation-dissipation theorem (FDT) relates the response to small perturbations to the relaxation of spontaneous fluctuations [18].

In general, we can measure the fluctuations and responses of any observable. Here, we focus on the energy of a given subunit. In both of our model systems, the total energy comes from interactions between pairs of subunits, $E_0=(1/2)\Sigma'_{ij}E_{ij}$, where the primed sum excludes terms with $i=j$. We denote the energy of the ith monomer by

$$E_i \equiv (1/2)\sum_{j(\neq i)} E_{ij}. \tag{1}$$

We measure the responses in the system as follows. Starting from a given initial state, the system assembles for a waiting time t_w. We then turn on a perturbation to the energy, $\delta E(t)=\Sigma_i h_i E_i \Theta(t-t_w)$, where h_i is the (small) field applied to the ith subunit, and $\Theta(x)$ is the unit-step function. We measure the integrated response to this field,

$$\chi(t,t_w)=\left(\frac{\partial\langle E_i(t)\rangle_{t_w}}{\partial(\beta h_i)}\right)_{\boldsymbol{h}=\boldsymbol{0}}, \tag{2}$$

where the notation $\boldsymbol{h}=(h_1,h_2,\ldots)$, and β^{-1} is the temperature multiplied by Boltzmann's constant. The average is over trajectories of the system in the presence of the perturbation. In practice, we evaluate the partial derivative by assigning $h_i=\delta h$ to half of the subunits (selected at random), and $h_i=-\delta h$ to the other half. In the linear response regime (small δh), the mean energy at $\boldsymbol{h}=\boldsymbol{0}$ can then be estimated by $\bar{E}(t)=N^{-1}\Sigma_i E_i(t)$, and the response by $\Sigma_i[E_i(t)-\bar{E}(t)]/h_i$. These quantities are then averaged over many independent runs of the dynamics. Our results for the capsid system were obtained at $\delta h=0.05$ and those for the disk system were obtained at $(\delta h/T)=0.3$. These values are small enough that our estimates of $\chi(t,t_w)$ change very little if δh is reduced, which indicates that we are in the linear response regime. For systems with Monte Carlo dynamics, such as the disk system, the derivative in Eq. (2) can also be evaluated as a correlation function for the unperturbed dynamics, in which case it is no longer necessary to apply the field h_i directly [19,20].

For a system at equilibrium, the fluctuation-dissipation theorem [18] states that

$$\chi(t,t_w)=C(t,t)-C(t,t_w) \tag{3}$$

for all t and t_w, where

$$C(t,t_w)\equiv\langle E_i(t)E_i(t_w)\rangle-\langle E_i(t)\rangle\langle E_i(t_w)\rangle. \tag{4}$$

Alternatively, we can define the response to an instantaneous perturbation (impulse response), as a derivative of the integrated response: $R(t,t_w)=-\partial\chi(t,t_w)/\partial t_w$. In that case, the FDT states that

$$R(t,t_w)=\frac{\partial C(t,t_w)}{\partial t_w}. \tag{5}$$

Away from equilibrium, we define the correlation-response ratio $X(t,t_w)$ by

$$R(t,t_w)=X(t,t_w)\frac{\partial C(t,t_w)}{\partial t_w}. \tag{6}$$

Thus, $X(t,t_w)$ is the response of the system to an instantaneous perturbation, normalized by the response of an equilibrium system with the same fluctuations.

In equilibrium, the fluctuation-dissipation theorem implies that $X(t,t_w)=1$ for all t and t_w. Away from equilibrium, $X(t,t_w)$ may take any value. It is most conveniently obtained from the gradient of a parametric plot of $\chi(t,t_w)$ against $C(t,t_w)$, where the parametric variable is the waiting time t_w [21]. We will see that parametric plots distinguish between systems above the kinetic crossover region, and those below it. This application of the FDR is the main result of this paper.

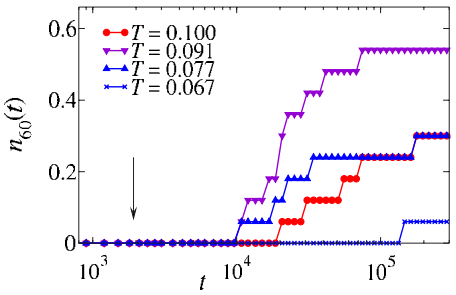

FIG. 3. (Color online) Sample trajectories in the capsid system, showing $n_{60}(t)$, defined as the fraction of particles in bonded clusters of size 60. We use a logarithmic scale for the reduced time t. The fraction $n_{60}(t)$ reflects the number of capsids in the system, since disordered clusters containing exactly 60 subunits are rare. The first capsids appear at times around 10^4. The system is away from global equilibrium until reduced times are at least of the order of 10^5. The arrow indicates the maximal time associated with our measurements of correlation and response functions (Figs. 4–6).

B. Results

In Fig. 3, we illustrate the time scale associated with capsid formation. The first capsids form in the system at times around 10^4, and all systems shown are significantly out of equilibrium until reduced times at least as large as 10^5. The yield measurements of Fig. 1 were taken at $t=3\times10^5$. As time proceeds, the system evolves increasingly slowly toward the equilibrium state. We will show correlation and response data at times of order 10^3, so the system is still well away from global equilibrium in all cases. However, we will find that systems at temperatures above the kinetic crossover region have responses in accordance with FDT, while those below it do not. In the disc system, the snapshots of Fig. 2 show that the system is well away from equilibrium at times around 5×10^6. For that system, we will show correlation-response data at much earlier times, those less than 10^5.

Some results for the capsid response function are shown in Fig. 4, where we show how multiple simulations are used to plot the response as a function of t_w for fixed t, which is useful for estimating the impulse response. A typical correlation function is shown in Fig. 5.

Results for the FDR in both capsid and disk systems are shown in Fig. 6, where we have normalized both correlation and response by the equal time fluctuation $C(t,t)$. [The function $C(t,t)$ is independent of t_w, so the gradient of the parametric plot is $-X(t,t_w)$, and is unaffected by the normalization.] Above the kinetic crossover, assembly is taking place, but the energy response is in accordance with the FDT, so $X(t,t_w)\approx1$, at least for the times that we considered. As the system passes through the kinetic crossover, the FDR shrinks. While it can be convenient to characterize this crossover by the single temperature T^*, it is more accurate to think of a temperature range over which the long-time behavior of system changes smoothly from effective to ineffective as-

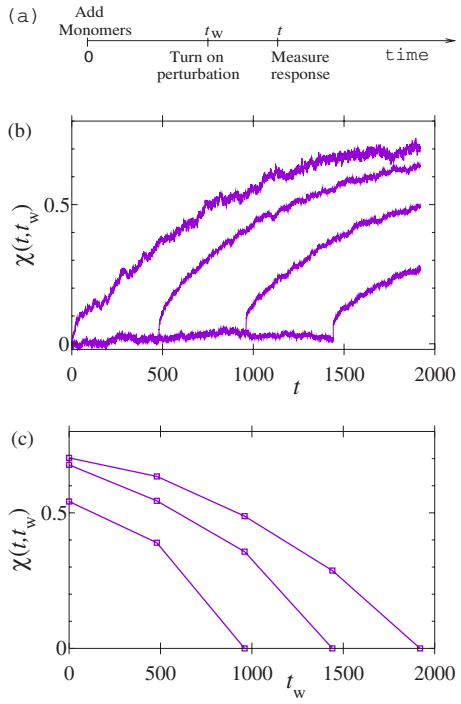

FIG. 4. (Color online) (a) A time line illustrating the simulation protocol used to measure the response. (b) Response in the capsid system (in units of ε_b^2) at reduced temperature $T=0.091$, as a function of time t, for $t_w=0,480,960,1440$. The data are plotted with lines, since each simulation yields data points for all t. (c) Plot of the response as a function of waiting time t_w, for $t=960,1440,1920$. This is a replot of some of the data of the middle panel, but it allows estimation of the impulse response $\partial\chi(t,t_w)/\partial t_w$. In this case, the data are shown as points (squares), and points with the same value of t are connected by lines.

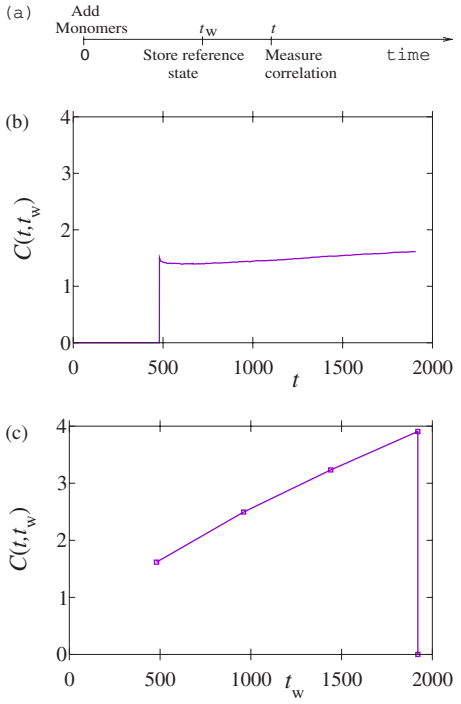

FIG. 5. (Color online) (a) A time line indicating the simulation protocol used to measure the correlation. (b) Correlation function in the capsid system (in units of ε_b^2) at $T=0.091$, as a function of reduced time t, for $t_w=480$. (c) Correlation as a function of t_w, for $t=1920$. The absence of time-reversal symmetry is clear.

sembly. This smooth change is accompanied by a smooth change in the FDR.

We conclude that, if a system is to be designed so that it assembles effectively, the correlation-response ratio can be used to obtain a general prediction for the regime of good assembly, before running the long simulations required to test the yield directly.

Finally, note that we constructed Fig. 6 using data at constant t and variable t_w, since the gradient of this plot gives the FDR. This procedure requires a separate simulation for each data point. However, if we only wish to know if the integrated response is small or large compared to the FDT prediction, it is sufficient to use data at a single t_w: a simple comparison of $C(t,t)-C(t,t_w)$ and $\chi(t,t_w)$ is already quite informative in that case (note, however, [21]).

IV. DISCUSSION

We now consider the kinetic and thermodynamic crossovers in a little more detail. We then discuss how the change in FDR at the kinetic crossover arises, and the extent to which we expect it to generalize to other self-assembling systems.

A. Thermodynamic and kinetic crossovers

We measure the yield of our assembly processes by running long simulations of length t_{yield} (recall Figs. 1 and 2). These simulations have three types of final state. At high temperatures, no assembly takes place, and the system consists primarily of free subunits. At low temperatures, the system evolves into a state that consists primarily of disordered metastable clusters.

We also find an intermediate temperature regime, in which the final state has a substantial quantity of assembled products. This regime is delineated by two crossovers. For an operational definition of the high-temperature crossover, we impose a threshold on the relative probabilities of bonded and free subunits at time t_{yield}. While this definition depends on t_{yield}, the position of the crossover has a well-defined limit as $t_{\text{yield}}\to\infty$, which can be evaluated from the contribution of free subunits to the thermodynamic partition function of the system. Thus we refer to this crossover as "thermodynamic."

To define the low-temperature crossover, we consider the relative probabilities of disordered clusters and correctly assembled products at t_{yield}. As the temperature is reduced, the maximum of the yield occurs when the disordered clusters are common enough to significantly impede assembly. We therefore identify this maximum with the low-temperature crossover. If we calculate the yield in the equilibrium state, we expect it to depend monotonically on the temperature, since the correctly assembled states minimize the total energy in both of our systems. Thus, the presence of the maximum in the yield is a kinetic effect that arises from the slow

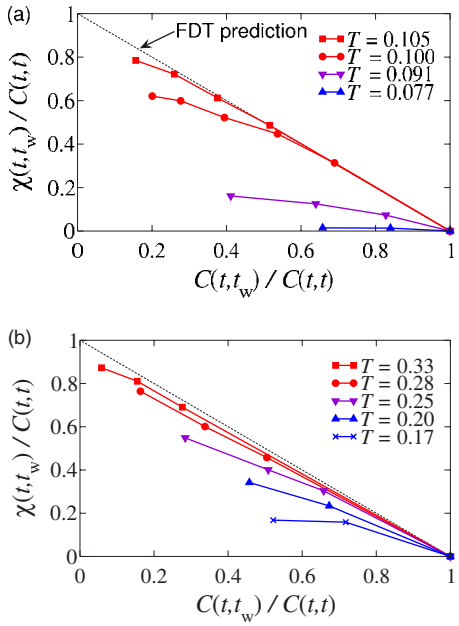

FIG. 6. (Color online) Correlation response plots for (a) the capsid system at $t=1920$ and $960<t_w<t$, and (b) the disk system at $t=8\times10^4$. These systems are all well away from equilibrium, but the response is in accordance with the prediction of the FDT at the higher temperatures. The response decreases rapidly as the system passes through the kinetic crossover and falls out of equilibrium. The red and blue coloring is consistent with that of Figs. 1 and 2.

annealing of disordered clusters. This motivates our use of the term "kinetic crossover." Clearly, the existence of a regime of efficient assembly requires that the kinetic crossover is not too close to the thermodynamic one. If the system crosses over smoothly from free subunits to disordered clusters, then there is no temperature at which assembly is efficient on the time scale t_{yield}.

B. Local equilibration

We now return to the link between the kinetic crossover and the FDR. The general idea is that dynamics that is locally time-reversal symmetric allows disordered states to anneal into ordered states. This idea is not new (for example, see Ref. [22], especially its Fig. 1). However, the FDR provides a quantitative measure of this effect.

The crystalline state of the two-dimensional system of disks is close packed, with each particle bonded to six neighbors. During assembly, the fraction of such particles in a given cluster provides a measure of its crystallinity. As clusters form, there are many possible states with low crystallinity, and the system tends to visit these states quite frequently. The effectiveness of assembly depends on whether these states are able to anneal into crystalline clusters. This annealing becomes more difficult as the disordered clusters grow. For example, annealing the disordered clusters of Fig. 2 into crystallites requires highly cooperative processes with large activation energies, while annealing small disordered clusters requires less cooperativity.

Our results indicate that, near optimal assembly, large disordered clusters are avoided because the system remains *locally equilibrated* at each stage of the assembly process (although the system is globally out of equilibrium). At any stage of assembly, there will be a set of likely states. The condition of local equilibration is that the relative probabilities of these likely states reflect their relative Boltzmann weights. If this condition holds, the system avoids the disordered states that are precursors to the large disordered clusters of Fig. 2. For example, small disordered clusters have smaller Boltzmann weights than crystalline clusters of the same size, so local equilibration suppresses the disordered states. On the other hand, if disordered states are likely at any stage of assembly, this indicates that they are not being annealed into crystallites, and are likely to evolve into larger disordered clusters.

To link this argument with the FDR, we first demonstrate a link between local equilibration and an approximate time-reversal symmetry. We consider two states $\mathcal{C}$ and $\mathcal{C}'$ that are both likely at a given stage of assembly. The rate with which the system makes transitions from $\mathcal{C}$ to $\mathcal{C}'$ is

$$\gamma(\mathcal{C}\to\mathcal{C}',t)=W(\mathcal{C}'|\mathcal{C})p(\mathcal{C},t), \tag{7}$$

where $p(\mathcal{C},t)$ is the probability that the system is in state $\mathcal{C}$ at time t, and $W(\mathcal{C}'|\mathcal{C})$ is the rate for transitions to state $\mathcal{C}'$ given that the system is initially in state $\mathcal{C}$. [The rate $W(\mathcal{C}|\mathcal{C}')$ depends only on the dynamical rules of the model, while the rate $\gamma(\mathcal{C}\to\mathcal{C}',t)$ depends also on the state of the system at time t.]

For models that obey detailed balance, we have

$$W(\mathcal{C}'|\mathcal{C})\exp(\beta E_{\mathcal{C}'})=W(\mathcal{C}|\mathcal{C}')\exp(\beta E_{\mathcal{C}}). \tag{8}$$

Further, if the system is locally equilibrated, then we have

$$p(\mathcal{C},t)\exp(\beta E_{\mathcal{C}})\approx p(\mathcal{C}',t)\exp(\beta E_{\mathcal{C}'}), \tag{9}$$

where $\mathcal{C}$ and $\mathcal{C}'$ are likely states at this time. Thus, the rates for forward and reverse transitions between these states are equal:

$$\gamma(\mathcal{C}\to\mathcal{C}',t)\approx\gamma(\mathcal{C}'\to\mathcal{C},t). \tag{10}$$

This relation is an approximate time-reversal symmetry of the locally equilibrated state, which holds on time scales for which the set of likely states is not changing significantly.

The extent to which this approximate time-reversal symmetry holds is correlated with the degree of local equilibration, and hence with the extent to which the system is discriminating between high-energy disordered states and low-energy ordered ones. By avoiding the high-energy disordered states, the locally equilibrated system tends to assemble effectively.

To link this local equilibration with the FDR, we show in the Appendix that, for systems obeying detailed balance, deviations from the FDT arise from differences between the probabilities of trajectories and their time-reversed counterparts, during the time between perturbation and measurement. The key result is (A5). We conclude that the FDR is a probe of the degree to which the system obeys time-reversal symmetry between times t_w and t, and hence of the degree of local equilibration.

Thus, our results for both capsid and disk systems (Fig. 6) are consistent both with our hypothesis that the system falls out of local equilibrium at the kinetic crossover, and with our interpretation of the FDR as a measure of local equilibration. The parametric plots summarize the important features of the correlation and response functions, in a single system-independent plot, in which time and energy scales are rescaled away. The qualitative similarities in the behavior of the FDR are all the more remarkable given the different dimensionalities of the two models that we consider, and the very different structures of their assembled states.

C. Generic and nongeneric features of the FDR

While the behavior of both capsid and disk systems are both consistent with our analysis above, there are important differences between the two panels of Fig. 6. In particular, at the peak of the assembly curve, the response in the disk system is larger than the corresponding response in the capsid system.

The reason for this difference can be inferred from the states shown in Figs. 1 and 2. In the disk system, the crystallinity of the product is rather low at all temperatures. Even small clusters typically explore many disordered states before they form locally crystalline structures. The system needs to be very close to local equilibrium in order to ensure that the ordered structures can be discriminated from the large number of disordered states. Thus, assembly is effective only when the FDR is close to unity. On the other hand, the directional interactions in the model capsid system impose quite stringent constraints on the local structure of the growing cluster. This reduces the possibility for stable disordered clusters, and discriminating between ordered and disordered states is easier. Thus, the system still assembles effectively even when deviations from local equilibrium are quite significant, and assembly is still effective even when deviations from FDT are quite large.

Taking account of these differences, we emphasize the main feature of Fig. 6: the FDR is large above the kinetic crossover, and small below it. We expect this behavior to be preserved as long as three conditions are met. First, the observables used to construct the FDR should couple to the processes by which metastable disordered states are annealed into ordered ones. For example, if we had measured the FDR in the capsid system using the capsomer positions in place of their energies, then diffusive processes would dominate both correlation and response functions, and this response is not sensitive to the extent to which the bonds in the system are locally equilibrated.

Second, we require that the assembled state of the system minimizes the free energy both globally and locally. Many biological systems are believed to have funnel-shaped energy landscapes consistent with this constraint [23]. The models presented in this paper also have this property. We believe that satisfying this constraint contributes quite generally to good assembly, and it is therefore practical to bear it in mind when designing self-assembling systems. Of course, systems that violate this constraint do exist. For example, in three dimensions, minimization of the free energy of small clusters of spherical particles leads to icosahedral structures [24], while the crystalline phase has a close-packed structure. It is therefore possible for these particles to assemble into icosahedral structures while always remaining locally equilibrated. The FDR would be close to unity, but the system would never visit the "correctly assembled" close-packed structure.

The third condition that is required to ensure usefulness of the parametric FDR plot concerns the time t used to construct it. The behavior of Fig. 6 depends weakly on the value of t, but changing its order of magnitude will lead to different behavior. In particular, at very low temperatures and for large t, the capsid system shows a FDR close to unity. This occurs because the system is locally equilibrated over a particular set of disordered states. However, in this case, the system would not be locally equilibrated while the disordered clusters were forming, so that FDR on that time scale would have been smaller than unity. In other words, detection of the relevant deviations from local equilibrium requires a measurement on the time scales during which those deviations occur.

These three conditions show that the application of the FDR to self-assembling systems requires some consideration of the relevant observables and time scales. However, for the systems studied in this paper, meeting these conditions does not require careful tuning of model parameters, but only the kind of qualitative analysis discussed in this section. This represents evidence in favor of the applicability of these methods to other self-assembling systems.

V. OUTLOOK

The arguments of Sec. IV seem general, and relatively independent of details of the system. Further tests of the links between efficient assembly, local equilibration, and FDRs would be valuable, especially if FDRs could be measured experimentally. In principle, FDRs can be obtained whenever conjugate correlation and response functions can be measured. Measuring fluctuations and responses of local quantities, such as the energy of a single subunit, requires a high degree of experimental control, but methods do exist in some systems. For example, Wang *et al.* [25] recently measured a FDR in a three-dimensional glassy colloidal system. The diffusive correlation function is conjugate to the response of a single particle to a force in that case. Applying similar methods to ordering processes of spheres or disks would be analogous to our studies of the sticky disk system.

Turning to biological systems, it would be possible to measure the degree of kinetic frustration in the folding of biomolecules, either computationally in more detailed capsid models, or in systems such as the trpzip peptide [26], or, experimentally, in RNA folding, by a generalization of the experiment of [27]. In this latter case, the conjugate variables of force and displacement are already measurable, although obtaining good statistics for the correlations and responses as a function of both t and t_w might be challenging. Results obtained in this way would complement information about the nonequilibrium dynamics obtained from analysis of the work distribution [28,29]. For example, the thermodynamic

definitions of reversible and irreversible work are linked to the idea that nonequilibrium processes can occur with or without local equilibration. By characterizing the extent to which particular degrees of freedom are locally equilibrated on particular time scales, FDRs provide another link between these thermodynamic ideas and the statistical mechanics of nonequilibrium trajectories.

ACKNOWLEDGMENTS

R.L.J. was funded initially by NSF Grant No. CHE-0543158 and later by the Office of Naval Research Grant No. N00014–07–1–689. M.F.H. was supported initially by NIJ Grant No. F32 GM073424-01, and later by the HHMI-NIBIB Interfaces Initiative grant to Brandeis University. DC was funded initially by NSF Grant No. CHE-0543158 and later by NSF Grant No. CHE-0626324.

APPENDIX: TIME REVERSIBILITY, AND THE FDR

In this appendix, we briefly consider a general stochastic system evolving between times t_i and t_f, and show how deviations from the predictions of the FDT come from trajectories (histories) which occur with probabilities that are different from those of their time-reversed counterparts.

Consider a stochastic system evolving between times t_i and t_f. The energy of a configuration $\mathcal{C}$ during this time period is given by $E(\mathcal{C})=E_0(\mathcal{C})-hA(\mathcal{C})$, where h is a field, A is an observable, and $A(\mathcal{C})$ is its value in configuration $\mathcal{C}$. The stochastic dynamics obey detailed balance with respect to the Boltzmann distribution $p_{\mathrm{eq}}(\mathcal{C}) \propto e^{-\beta E(\mathcal{C})}$. The response of observable B to the field h is

$$\chi(t_f,t_i) = \sum_{\mathcal{C}(t)} B(\mathcal{C}_f) \left. \frac{\partial P[\mathcal{C}(t);h]}{\partial(\beta h)} \right|_{h=0}, \tag{A1}$$

where the sum is over trajectories (histories) of the system, which we indicate by the function $\mathcal{C}(t)$; the initial and final configurations of the trajectory are $\mathcal{C}_i$ and $\mathcal{C}_f$, respectively; and $P[\mathcal{C}(t);h]$ is the probability of a trajectory, which includes the probability of its initial condition.

The property of detailed balance dictates that

$$\log \frac{P[\mathcal{C}(t);h]}{p_i(\mathcal{C}_i)} - \log \frac{P[\bar{\mathcal{C}}(t);h]}{p_i(\bar{\mathcal{C}}_i)} = \beta h[A(\mathcal{C}_f) - A(\mathcal{C}_i)] - \beta[E(\mathcal{C}_f) - E(\mathcal{C}_i)], \tag{A2}$$

where $p_i(\mathcal{C}_i)$ is the probability of the initial condition of the trajectory $\mathcal{C}(t)$, and $\bar{\mathcal{C}}(t)$ is the time-reversed counterpart of $\mathcal{C}(t)$. That is, $\bar{\mathcal{C}}(t)=\mathbb{T}\mathcal{C}(t_i+t_f-t)$, where the operator $\mathbb{T}$ reverses all quantities that are odd under time reversal, such as momenta. To enforce time-reversal symmetry of the equilibrium state, we assume that the energy and its perturbation are time-reversal symmetric: $E(\mathcal{C})=E(\mathbb{T}\mathcal{C})$ and $A(\mathcal{C})=A(\mathbb{T}\mathcal{C})$. We also assume that $B(\mathcal{C})=B(\mathbb{T}\mathcal{C})$ for convenience, although the same analysis can also be carried out without this assumption, leading to an analogous result.

Using $(\partial P[\mathcal{C}(t);h]/\partial h)=P[\mathcal{C}(t);h](\partial/\partial h)\log P[\mathcal{C}(t);h]$, we substitute (A2) into (A1), and obtain

$$\chi(t_f,t_i) = \langle B(t_f)[A(t_f)-A(t_i)]\rangle + \sum_{\mathcal{C}(t)} B(\mathcal{C}_f)P[\mathcal{C}(t);0] \times \left. \frac{\partial}{\partial(\beta h)} \log P[\bar{\mathcal{C}}(t);h] \right|_{h=0}, \tag{A3}$$

where we have used $\langle\cdot\rangle \equiv \Sigma_{\mathcal{C}(t)}P[\mathcal{C}(t)](\cdot)$. The fluctuation-dissipation theorem states that the first two terms are equal at equilibrium, so we define $\Delta\chi(t_f,t_i)=\chi(t_f,t_i)-\langle B(t_f)[A(t_f)-A(t_i)]\rangle$ in order to measure deviations from the FDT.

To obtain an informative expression for $\Delta\chi(t_f,t_i)$, we use conservation of probability to write

$$\sum_{\mathcal{C}(t)} B(\bar{\mathcal{C}}_i)P[\bar{\mathcal{C}}(t);h] = \sum_{\mathcal{C}} B(\mathcal{C})p_i(\mathcal{C}). \tag{A4}$$

Thus, the derivative of the left-hand side of (A4) with respect to h is zero. Noting that $B(\bar{\mathcal{C}}_i)=B(\mathcal{C}_f)$, we subtract this derivative from the right-hand side of (A3), arriving at

$$\Delta\chi(t_f,t_i) = \sum_{\mathcal{C}(t)} B(\mathcal{C}_f) \left. \frac{\partial}{\partial(\beta h)} \log P[\bar{\mathcal{C}}(t);h] \right|_{h=0} \times \{P[\mathcal{C}(t);0] - P[\bar{\mathcal{C}}(t);0]\}. \tag{A5}$$

The purpose of (A5) is to show that, if all trajectories $\mathcal{C}(t)$ have the same probabilities as their time-reversed counterparts $\bar{\mathcal{C}}(t)$, then the second term vanishes, and the FDT applies. This condition holds exactly only at equilibrium, but if the dynamics of the system are close to local equilibrium between times t_i and t_f, then the relative weights of forward and reverse trajectories will be similar, and deviations from FDT will be small.

[1] T. D. Pollard and G. G. Borisy, Cell **112**, 453 (2003); J. B. Moseley and B. L. Goode, Microbiol. Mol. Biol. Rev. **70**, 605 (2006).

[2] S. Inoue and E. D. Salmon, Mol. Biol. Cell **6**, 1619 (1995); A. Desai and T. J. Mitchison, Annu. Rev. Cell Dev. Biol. **13**, 83 (1997).

[3] B. Antonny, P. Gounon, R. Schekman, and L. Orci, EMBO Rep. **4**, 419 (2003); K. Matsuoka, R. Schekman, L. Orci, and J. E. Heuser, Proc. Natl. Acad. Sci. U.S.A. **98**, 13705 (2001); S. C. Harrison and T. Kirchhausen, Cell **33**, 650 (1983).

[4] H. Fraenkel-Conrat and R. C. Williams, Proc. Natl. Acad. Sci. U.S.A. **41**, 690 (1955); A. Klug, Philos. Trans. R. Soc. Lon-

don, Ser. B **354**, 531 (1999); A. Zlotnick *et al.*, Virology **277**, 450 (2000); J. Sun *et al.*, Proc. Natl. Acad. Sci. U.S.A. **104**, 1354 (2007).

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

[6] S. C. Glotzer, Science **306**, 419 (2004); G. M. Whitesides and B. Grzybowski, *ibid.* **295**, 2418 (2002); T. Douglas and M. Young, *ibid.* **312**, 873 (2006).

[7] C. Valery *et al.*, Proc. Natl. Acad. Sci. U.S.A. **100**, 10258 (2003); H. Yan *et al.*, Science **301**, 1882 (2003); E. Strable *et al.*, Nano Lett. **4**, 1385 (2004); A. S. Blum *et al.*, *ibid.* **4**, 867 (2004); S. H. Park *et al.*, *ibid.* **5**, 729 (2005); S. K. Dixit *et al.*, *ibid.* **6**, 1993 (2006); P. O'Neill *et al.*, *ibid.* **6**, 1379 (2006).

[8] See, for example, F. H. C. Crick and J. D. Watson, Nature (London) **177**, 473 (1956); D. L. Caspar and A. Klug, Cold Spring Harbor Symp. Quant. Biol. **27**, 1 (1962); R. Zandi *et al.*, Proc. Natl. Acad. Sci. U.S.A. **101**, 15556 (2004).

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

[10] A. Louis *et al.*, e-print arXiv:cond-mat/0606634.

[11] See, for example, L. F. Cugliandolo and J. Kurchan, Phys. Rev. Lett. **71**, 173 (1993); J. Phys. A **27**, 5749 (1994); J. Kurchan and L. Laloux, *ibid.* **29**, 1929 (1996); S. Franz, M. Mezard, G. Parisi, and L. Peliti, Phys. Rev. Lett. **81**, 1758 (1998); A. Crisanti and F. Ritort, J. Phys. A **36**, R181 (2003); J. Kurchan, Nature (London) **433**, 222 (2005).

[12] R. L. Jack, L. Berthier, and J. P. Garrahan, J. Stat. Mech.: Theory Exp. (2006) P12005.

[13] S. Fielding and P. Sollich, Phys. Rev. Lett. **88**, 050603 (2002); P. Mayer, L. Berthier, J. P. Garrahan, and P. Sollich, Phys. Rev. E **68**, 016116 (2003); *ibid.* **70**, 018102 (2004).

[14] P. Mayer, S. Leonard, L. Berthier, J. P. Garrahan, and P. Sollich, Phys. Rev. Lett. **96**, 030602 (2006); S. Leonard, P. Mayer, P. Sollich, L. Berthier, J. P. Garrahan, e-print arXiv:cond-mat/0703164.

[15] S. Whitelam and P Geissler, e-print arXiv:cond-mat/0508100.

[16] R. H. Swendsen and J.-S. Wang, Phys. Rev. Lett. **58**, 86 (1987).

[17] P. Meakin, Phys. Rev. Lett. **51**, 1119 (1983); M. Kolb, R. Botet, and R. Jullien, *ibid.* **51**, 1123 (1983).

[18] D. Chandler, *Introduction to Modern Statistical Mechanics* (Oxford University Press, New York, 1987).

[19] L. Berthier, Phys. Rev. Lett. **98**, 220601 (2007).

[20] C. Chatelain, J. Stat. Mech.: Theory Exp. (2004) P06006.

[21] Defining analogous quantities to the FDR, such as the ratio of derivatives of correlation and response with respect to the final time t, is also possible. However, the FDR is the most physical and informative quantity. See, for example, Ref. [14], and Fig. 2 of P. Mayer and P. Sollich, J. Phys. A **37**, 9 (2004).

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

[23] A. Sali, E. I. Shakhnovich, and M. Karplus, Nature (London) **369**, 248 (1994); J. D. Bryngelson, J. N. Onuchic, N. D. Socci, and P. G. Wolynes, Proteins **21**, 167 (1995); A. Zlotnick, J. Mol. Recognit. **18**, 479 (2005).

[24] H. Jonsson and H. C. Andersen, Phys. Rev. Lett. **60**, 2295 (1988); J. P. K. Doye, M. A. Miller, and D. J. Wales, J. Chem. Phys. **110**, 6896 (1999).

[25] P. Wang, C. Song, and H. A. Makse, Nat. Phys. **2**, 526 (2006).

[26] W. Y. Yang *et al.*, J. Mol. Biol. **336**, 241 (2004).

[27] B. Onoa *et al.*, Science **299**, 1892 (2003).

[28] J. Liphardt *et al.*, Science **296**, 1832 (2002).

[29] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997); G. E. Crooks, Phys. Rev. E **61**, 2361 (2000); F. Ritort, in *Seminaire Poincaré 2003*, edited by J. Daibard, B. Duplantier, and V. Rivasseau (Birkhauser, Basel, 2004).