Understanding the Concentration Dependence of Viral Capsid Assembly Kinetics—the Origin of the Lag Time and Identifying the Critical Nucleus Size

Michael F. Hagan* and Oren M. Elrad
Department of Physics, Brandeis University, Waltham, Massachusetts

ABSTRACT The kinetics for the assembly of viral proteins into a population of capsids can be measured in vitro with size exclusion chromatography or dynamic light scattering, but extracting mechanistic information from these studies is challenging. For example, it is not straightforward to determine the critical nucleus size or the elongation time (the time required for a nucleus to grow to completion). In this work, we study theoretical and computational models for capsid assembly to show that the critical nucleus size can be determined from the concentration dependence of the assembly half-life and that the elongation time is revealed by the length of the lag phase. Furthermore, we find that the system becomes kinetically trapped when nucleation becomes fast compared to elongation. Implications of this constraint for determining elongation mechanisms from experimental assembly data are discussed.

INTRODUCTION

The assembly of protein building blocks into a shell or capsid is essential for viral replication and thus understanding the mechanisms by which assembly proceeds could identify targets or opportunities for novel antiviral therapies. However, despite extraordinary progress in determining the structures of assembled capsids, assembly mechanisms for most viruses remain poorly understood because the structures of transient assembly intermediates have been inaccessible experimentally. The kinetics for spontaneous capsid assembly in vitro have been measured with size exclusion chromatography (SEC) and x-ray and light scattering (e.g., (1–7)), but extracting mechanistic information such as the critical nucleus size or the time to assemble an individual capsid has been challenging. In this article, we theoretically examine two models for capsid assembly kinetics to show that these properties can be determined from the concentration dependence of median assembly times and the lag phase.

Assembly kinetics in vitro has been measured for a number of icosahedral viruses (e.g., (1–7)) and demonstrates sigmoidal growth characterized by a lag phase, rapid growth, and finally saturation (see Figs. 1 and 2). Zlotnick and coworkers (2,8,9) showed that partial capsid intermediates assemble during the lag phase, but it has often been assumed that the duration of the lag phase corresponds to the time required for the concentration of critical nuclei to reach steady state, in analogy to models of actin nucleation. However, in this work we show that because light-scatter signal measures the mass-averaged molecular weight of assemblages (4) and SEC usually monitors complete capsids, the length of the lag phase is related to the elongation time, or the time required for a nucleated partial capsid to grow to completion. Similarly, the critical nucleus size cannot be reliably determined from the concentration dependence of initial or maximum growth rates (9), and a method to do so using the extent of assembly (9) is data-intensive. We demonstrate that the critical nucleus size can be identified in a straightforward manner from the concentration dependence of the median assembly time. Finally, we show that the system becomes kinetically trapped when the elongation time becomes long compared to the timescale for nucleation. It is important to note that we do not consider the effects of a slow transition between assembly-active and assembly-incompetent conformations of free subunits, recently suggested in Chen et al. (6).

While preparing this article, we became aware of a related study in which Morozov et al. (10) consider simplified capsid assembly models in which nucleation occurs via a single dimerization event, which enables a elegant analytic solution. They show that the early phase of assembly can be characterized as a shock front, and that for some conditions prohibitively long timescales are required to reach equilibrium. In this work, we consider nucleation as a multistep subunit addition process, with the objectives of understanding the concentration dependence of overall assembly times and inferring nucleation and capsid growth times from experimental light scatter measurements.

This article is organized as follows. In the next section, we use scaling laws for elongation and nucleation times as a function of subunit concentration in the context of a simplified assembly reaction. We then demonstrate that these estimates hold for a more realistic model by examining Brownian dynamics simulations for a model of assembly into icosahedral shells. We then further explore assembly timescales using a system of rate equations for the assembly of polyhedral shells and two models for capsid intermediate free energies. In Discussion and Outlook, we consider the

Submitted April 29, 2009, and accepted for publication November 18, 2009.

*Correspondence: hagan@brandeis.edu

Editor: Gregory A. Voth.

0006-3495/10/03/1065/10 $2.00

doi: 10.1016/j.bpj.2009.11.023

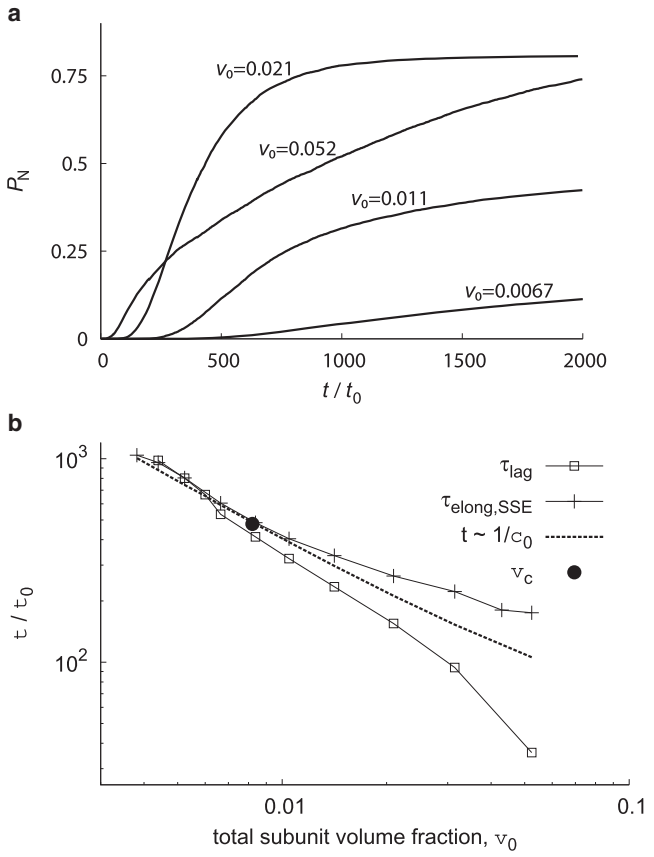

FIGURE 1 (*a*) The completion fraction P_N in Brownian dynamics (BD) simulations (the NVT simulations) is shown as a function of time for indicated initial subunit volume fractions (v_0). (*b*) Lag times τ_{lag}, elongation times τ_{elong}, and nucleation times τ_{nuc}^{min} calculated from BD simulations are shown as functions of the initial subunit volume fraction. Lag times are calculated from canonical ensemble simulations as described in Fig. 3, whereas τ_{elong} is calculated from the steady-state ensemble as described in the text. The volume fraction v_c corresponding to the crossover concentration is indicated by the •. The dashed line is a guide to the eye to demonstrate scaling inverse to subunit concentration. The scatter in the points at volume fractions below v_c gives an indication of the statistical error at low concentrations.

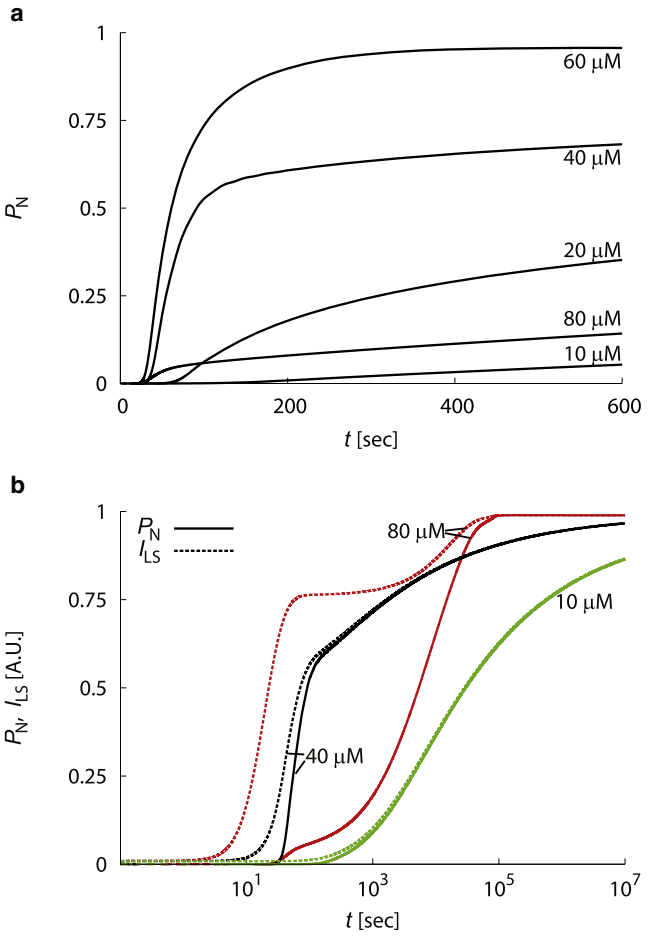

FIGURE 2 The time dependence of capsid assembly for the NG model varies with initial subunit concentration c_0. (*a*) The completion fraction P_N as a function of time for indicated initial subunit concentrations. (*b*) The calculated light scatter closely tracks completion fraction until kinetic trapping sets in. The calculated light scatter (*dashed lines*) and completion fraction (*solid lines*) are shown as a functions of time (on a logarithmic scale) for indicated initial subunit concentrations, with $g_{nuc} = -7k_BT$, $N = 120$, and $f = 10^5$ M^{-1}s^{-1}.

implications of our findings for experimental measurements of capsid assembly dynamics. Additional simulation details, derivations, and further evaluation of approximations in the rate equation models are given in the Supporting Material.

CAPSID ASSEMBLY TIMESCALES

In this section, we develop scaling laws for the concentration dependence of the duration of lag phase and nucleation timescales. To facilitate the presentation, we first consider a systematic assembly process which is greatly simplified in comparison to capsid assembly. In the following sections and in the Supporting Material, however, we show that the conclusions of this section remain valid when the simplifications are eliminated.

We consider a system of capsid protein subunits with total concentration c_0 that start assembling at the time $t = 0$ into capsids; the word subunit refers to the basic assembly unit, which could be a protein dimer or capsomer. Our simplified reaction is given by

$$1 \underset{b_{nuc}}{\overset{fc_1}{\rightleftharpoons}} 2 \underset{fc_1}{\overset{b_{nuc}}{\rightleftharpoons}} \cdots \underset{b_{nuc}}{\overset{fc_1}{\rightleftharpoons}} n_{nuc} \underset{b_{elong}}{\overset{fc_1}{\rightleftharpoons}} \cdots \underset{b_{elong}}{\overset{fc_1}{\rightleftharpoons}} N, \qquad (1)$$

where N is the number of subunits in a capsid, c_1 is concentration of unassembled subunits, and b_i is the dissociation rate constant (with $i = \{\text{nuc, elong}\}$), which is related to the forward rate constant by the equilibrium constant, $b_i = f\exp(g_i/k_BT)/v_0$, with g_i the subunit association free energy (arising from hydrophobic and electrostatic interactions (11,12)) and v_0 the standard state volume. The nucleation and elongation phases are distinguished by the fact that

association in the nucleation phase is not free-energetically favorable, $c_1 \exp(-g_{\rm nuc}/k_{\rm B}T) < 1$, whereas association in the elongation phase is favorable, $c_1 \exp(-g_{\rm elong}/k_{\rm B}T) > 1$. For the moment, we assume that there is an average nucleus size $n_{\rm nuc}$; identifying $n_{\rm nuc}$ from assembly kinetics data is one of the objectives of our work. The physical origins of nucleation and the factors that determine $n_{\rm nuc}$ are discussed in Rate Equation Models for Capsid Assembly.

We write the overall capsid assembly time τ as $\tau = \tau_{\rm nuc} + \tau_{\rm elong}$, with $\tau_{\rm nuc}$ and $\tau_{\rm elong}$ the average times for nucleation and elongation, respectively. For all of the models considered in this work, we will see that when elongation is fast compared to nucleation, the duration of the lag phase is given by the mean elongation time for the first capsids to assemble: $\tau_{\rm lag} = \tau_{\rm elong}(t = 0)$. Thus, for this model the lag time can be calculated from the mean first-passage time for a biased random walk with a reflecting boundary conditions at $n_{\rm nuc}$ and absorbing boundary conditions at N, with forward and reverse hopping rates given by fc_0 and $b_{\rm elong} = f\exp(g_{\rm elong})/v_0$, respectively. Mean first-passage times for these boundary conditions are derived in Bar-Haim and Klafter (13); inserting these forward and reverse hopping rates yields

$$\tau_{\rm elong} = \frac{n_{\rm elong}}{fc_0 - b_{\rm elong}} - \left(\frac{b_{\rm elong}}{fc_0 - b_{\rm elong}}\right)^2 \left(\frac{b_{\rm elong}}{fc_0}\right)^{n_{\rm elong}}, \quad (2)$$

with $n_{\rm elong} = N - n_{\rm nuc}$.

In the limit of $fc_0 \gg b_{\rm elong}$, Eq. 2 can be approximated to give $t_{\rm elong} \approx n_{\rm elong}/fc_0$, whereas for similar forward and reverse reaction rates, $fc_0 \approx b_{\rm elong}$, it approaches the solution for an unbiased random walk $t_{\rm elong} \approx n_{\rm elong}^2/2fc_0$. We calculate elongation times for other models in the Appendix and measure them for a more realistic model in the next section.

Under conditions of constant free subunit concentration, we can derive the average nucleation time with an equation analogous to Eq. 2 (9,14),

$$\tau_{\rm nuc}^{\rm min} = \frac{\hat{n}}{fc_0 - b_{\rm nuc}} - \left(\frac{b_{\rm nuc}}{fc_0 - b_{\rm nuc}}\right)^2 \left(\frac{b_{\rm nuc}}{fc_0}\right)^{\hat{n}} \approx f^{-1}\exp(G_{\hat{n}}/k_{\rm B}T)c_0^{-\hat{n}}, \quad (3)$$

with $\hat{n} = n_{\rm nuc} - 1$, and $G_{\hat{n}}$ the interaction free energy for the pre-nucleus ($G_{\hat{n}} = [\hat{n} - 1]g_{\rm nuc}$ for this model). However, because free subunits are depleted by assembly, the net nucleation rate never reaches this value and asymptotically approaches zero as the reaction approaches equilibrium. Instead, treating the system as a two-state reaction with $n_{\rm nuc}$-th order kinetics (see the Supporting Material) yields an approximation for the median assembly time $\tau_{1/2}$, the time at which the reaction is 50% complete

$$\tau_{1/2} \approx \frac{2^{\hat{n}} - 1}{\hat{n}} \frac{P_{\rm N}^{\rm eq}}{Nf} \exp(G_{\hat{n}}/k_{\rm B}T)c_0^{-\hat{n}}, \quad (4)$$

with $P_{\rm N}^{\rm eq}$ as the equilibrium fraction of subunits in complete capsids, which can be measured experimentally (11). The factor of N^{-1} in Eq. 4 accounts for the fact that N subunits are depleted by each assembled capsid.

For all models considered, we will see that when capsid growth times are negligible compared to nucleation times, $\tau_{\rm elong}$ and Eq. 4, respectively, predict the duration of the lag phase and the overall median assembly time. However, as first noted by Zlotnick (8), the reaction becomes kinetically trapped if free subunits are depleted before most capsids finish assembling. It was recently suggested (10,14) that this trap occurs at binding free energies G_n and subunit concentrations c_0 for which the rate of subunit depletion by nucleation ($N/\tau_{\rm nuc}^{\rm min}$) is equal to the elongation rate. We find that the relationships between $\tau_{\rm elong}$ and assembly times begin to fail at a crossover concentration $c_{\rm c}$ for which initial nucleation and elongation rates are equal, but the system becomes kinetically trapped at a larger concentration $c_{\rm kt}$ defined by the point at which the median assembly time $\tau_{1/2}$ matches the elongation time. These concentrations are related to binding free energies and other parameters by

$$\tau_{\rm elong} \approx \tau_{\rm nuc}^{\rm min}/N \quad \text{for} \quad c_0 = c_{\rm c}$$
$$\tau_{\rm elong} \approx \tau_{1/2} \quad \text{for} \quad c_0 = c_{\rm kt}, \quad (5)$$

with $\tau_{\rm nuc}^{\rm min}$ and $\tau_{1/2}$, respectively, given by Eqs. 3 and 4.

INVESTIGATING THE LAG PHASE WITH BROWNIAN DYNAMICS (BD) SIMULATIONS

There are a number of simplifications in the schematic assembly process of the previous section:

- Malformed capsids are not considered (15–22);
- Assembly proceeds along a single pathway (23–25);
- Only single subunits can bind or unbind;
- Transitions between intermediates are allowed through binding or unbinding of a single subunit; and
- There is only one (average) forward rate constant f.

In this section, we show that our results are valid when applied to a computational model that does not make any of those simplifications. We specifically consider the conclusions about lag times because calculating median assembly times is computationally demanding at low concentrations; we examine nucleation times in the next section.

There are two observations about lag times made in this work to be checked:

- The first, crucial, observation is that lag times correspond to the mean of the distribution of initial capsid elongation times below $c_{\rm c}$.
- The second observation is that the mean elongation time varies inversely with free subunit concentration if elongation is primarily a first-order reaction; note that lag times will still correspond to mean elongation times even if elongation is not first-order.

We consider simulations of a model for the assembly of icosahedral shells (14,16,20), in which subunits are modeled as rigid bodies, where excluded volume interactions are modeled by spherically symmetric repulsive forces, and complementary subunit-subunit interactions that drive assembly are modeled by directional attractions. The lowest energy states in the model correspond to capsids, which consist of multiples of 30 subunits (each of which represents a protein dimer) in a shell with icosahedral symmetry. Because the spatial positions and orientations of all subunits are explicitly tracked, there are no assumptions about assembly pathways or the structures that emerge from assembly.

Simulation parameters

The parameters of the model are the energy associated with the attractive potential, ε_b, and the specificity of the directional attractions, which is controlled by the angular parameters θ_m and ϕ_m. Subunit positions and orientations are propagated according to overdamped Brownian dynamics, with the unit of time $t_0 = a^2/D$, where D is the subunit diffusion coefficient and a is the subunit diameter. Full details of the model are given in Hagan (14).We simulated systems with 2000 subunits in periodic boxes with side lengths ranging from 27 to 65, where all distances are measured in units of the subunit diameter a. These side lengths correspond to subunit volume fractions of $v_0 \in [0.052, 0.0038]$, corresponding to concentrations of $c_0 = 84$ μM to 1 mM, with $c_0 = 6v_0/(\pi a^3 N_A)$ with the subunit diameter $a = 5.2$ nm and N_A Avogadro's number. The interaction parameters were $\varepsilon_b = 12.25\ k_BT$, $\theta_m = 0.75$, and $\phi_m = \pi$. We consider the assembly of $T = 1$ capsids, so $N = 30$ dimer subunits.

NVT simulations

We consider two sets of simulations to evaluate the time dependence of capsid assembly. The first set corresponds to the in vitro empty-capsid experiments modeled throughout this work, and simulates capsid formation in the canonical (NVT) ensemble. Simulations are initialized by generating random positions and orientations for subunits, with subunit positions that lead to subunit-subunit overlap (positive potential energies in excess of 1 k_BT) rejected, and dynamics are integrated until a prescribed time.

The fraction of subunits in complete capsids P_N, which can be monitored by SEC, is shown as a function of time for several total subunit volume fractions in Fig. 1. In all cases there is a lag time, followed by the rapid appearance of complete capsids and then eventually saturation of growth. The rate of capsid formation is nonmonotonic with respect to initial subunit concentration; the highest subunit volume fraction ($v_0 = 0.052$), which is above c_c (Eq. 5), has the shortest lag time but a slow rise to saturation because the system becomes starved for free subunits or small oligomers and large oligomers rarely have geometries compatible with binding. Lag times are measured from simulation data using the procedure described in Fig. 3. Except at high concentrations, lag times roughly scale as $\tau_{lag} \propto v_0^{-1}$—indicative of a primarily first-order elongation reaction for these parameters, despite the fact that oligomer binding does occur for this model.

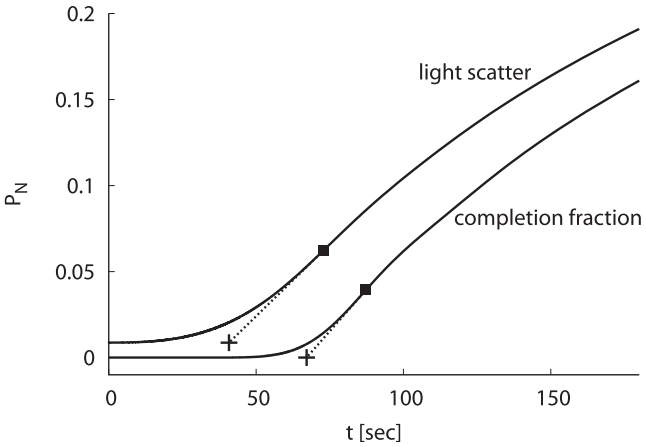

FIGURE 3 The end of the lag phase is measured by making a linear fit to the assembly kinetics trace at the point of maximal growth rate (■). The lag time (+) then corresponds to time at which the fit (*dashed line*) crosses the baseline. Plots are shown for the NG model with $c_0 = 20$ μM.

Steady-state ensemble simulations

To evaluate the correspondence between lag times and initial elongation times, it is necessary to directly measure mean elongation times of growing capsids in the simulations. However, achieving statistical relevance and avoiding finite size effects for the measurement of initial elongation times is extremely challenging in the NVT simulations because, by definition, most capsids assemble later in the simulation when free subunits have been depleted. Therefore, we also consider simulations in a steady-state ensemble (26), in which the free subunit concentration becomes time-independent. Specifically, clusters that become complete capsids, defined as closed shells in which each subunit has its full complement of four bonds, and clusters that reach a size of 35 or more subunits, are removed from the system and their subunits are reinserted into the simulation box, subject to the same overlap criterion as the initial state. The simulations reach steady state in a time that closely corresponds to the lag time measured in the NVT simulations. Once steady state is achieved, we measure the distribution of elongation times and the average nucleation rate by tracking individual clusters. These simulations were run for the same sets of parameters as the NVT simulations.

At steady state in the steady-state ensemble simulations, the concentration of free subunits c_1^{ss} is equal to the total subunit concentration c_0 minus the concentration of subunits in partially assembled capsids, which is dictated by the ratio

of capsid production rates and free subunit consumption rates. Below c_c, when elongation is fast compared to nucleation, the majority of subunits are free (or prenucleated) and c_1^{ss} is only slightly smaller than c_0, but the ratio drops significantly with increasing c_0, as shown in Fig. S7 in the Supporting Material.

The mean elongation times are shown as a function of the initial subunit volume fraction in Fig. 1. We see that for $v_0 \leq 0.006$ lag times measured from the NVT simulations and the mean elongation times in the steady-state ensemble agree to within error. Because the steady-state free subunit concentration c_1^{ss} closely corresponds to the initial free subunit concentration under these conditions (Fig. S7 in the Supporting Material), this finding supports the suggestion that lag times correspond to the initial mean elongation time. As the total subunit concentration increases above c_c, the steady-state concentration becomes smaller than the initial concentration $c_1^{ss} < c_0$ and thus the steady-state mean elongation times become longer than the lag times. We note, however, that above c_c lag times increase faster than $\tau_{lag} \propto 1/c_0$; similarly mean elongation times increase faster than $\tau_{elong} \propto 1/c_1^{ss}$. These results in part reflect the fact that binding of partial capsid intermediates becomes common above c_c (see below). Furthermore, above c_c, nucleation is no longer the rate-limiting step and lag times therefore correspond to the fastest members of the elongation time distribution, rather than the mean elongation time (i.e., when nucleation is rate-limiting, the first capsids to assemble have average elongation times, whereas those capsids with the shortest elongation times are the first to assemble when elongation is rate-limiting).

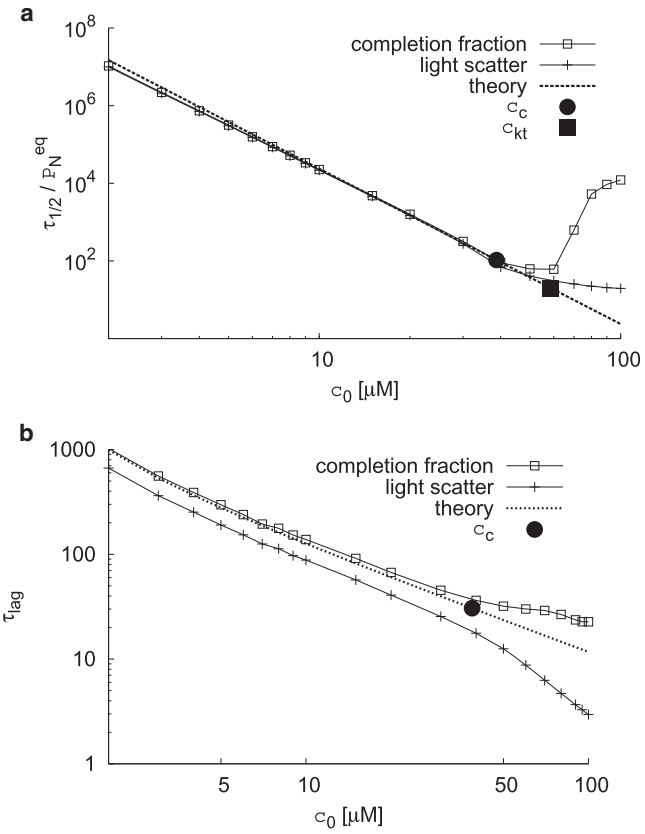

FIGURE 4 The median assembly times $\tau_{1/2}$ (*a*) and the lag times τ_{lag} (*b*) calculated from the completion fraction (P_N) and calculated light scatter (I_{LS}) are shown as functions of initial subunit concentration c_0. The estimates for the nucleation time (Eq. 4 with $n_{nuc} = 5$) and lag time Eq. 2 are shown as dashed lines, and the estimates for the crossover concentration c_c and kinetic trap concentration c_{kt} (Eq. 5) are shown as symbols on the estimated nucleation curve.

RATE EQUATION MODELS FOR CAPSID ASSEMBLY

Because evaluating median assembly times is computationally demanding at low concentrations (see Fig. 4 *a*), we use rate equation models to explore the relationship between nucleation times and overall capsid assembly timescales. Zlotnick and co-workers (2,8,9) have developed a system of rate equations that describe the time evolution of concentrations of empty capsid intermediates as

$$\begin{aligned}\frac{dc_1}{dt} &= -2f_1c_1^2 + b_2c_2 + \sum_{n=2}^{N} -f_nc_nc_1 + b_nc_n \\ \frac{dc_n}{dt} &= f_{n-1}c_1c_{n-1} - f_nc_1c_n \quad n = 2\ldots N \\ &\quad -b_nc_n + b_{n+1}c_{n+1},\end{aligned} \tag{6}$$

where c_n is the concentration of intermediates with n subunits, and f_n and b_n are, respectively, association and dissociation rate constants for intermediate n. There are several important assumptions, which are not present in the BD simulations: malformed capsids are not considered, only single subunits can bind or unbind, and only one f_i and b_i value is considered for each size i (averaged over averaged over all intermediates of that size). Despite these simplifications, rate equations of this form have shown good agreement with median assembly times of experimental assembly kinetics data (2,5). These simplifications should not affect our analysis of nucleation times, which involve a relatively small number of subunits and largely depend on the Boltzmann weights of pre-nuclei. Furthermore, we show that our results still hold when these simplifications are systematically eliminated in the Supporting Material.

The association and dissociation rate constants are related by detailed balance $b_n = f\exp(g_n/k_BT)/v_0$, with $g_n = G_n - G_{n-1}$ the change in free energy due to association of a subunit. Association free energies g_n, which can be fit to experimental data using the law of mass action (11,27,28), include hydrophobic and electrostatic interactions (12) and depend on pH and salt concentration (11). Specifying the assembly model requires defining the set of intermediates n and the transition rates f_n, b_n. To show that our conclusions are general, we consider two choices of these definitions.

The nucleation and growth (NG) model

Our first definition is based on the models of Zlotnick and co-workers (2,8,9), in which the subunit-subunit association free energy for intermediate i is proportional to the number of new subunit-subunit contacts $n_{c,i}$ formed by addition of a subunit to that intermediate. Note that this assumption neglects the fact that rotational entropy penalties are most likely not proportional to the number of contacts (see (35) for further discussion). Specifying $\{g_i\}$ thus requires defining the geometry of each intermediate. This usually begins with specifying the geometry of a capsid and its subunits in terms of a polyhedron (for example, see Fig. S6 in the Supporting Material or Fig. 1 in (9)), and assuming that assembly proceeds along a single path. The path can be comprised of the lowest energy intermediate for each size i (8) or correspond to an average pathway (9) in which all subunits, except during the initial and final stages of assembly, make the same average number of contacts. We choose the latter approach, as it is simpler and not obviously inferior in light of the other approximations in the rate equation models. Specifically, the association rate constant f is independent of intermediate size, and association free energies are given by $g_n = g_{\rm nuc}$ before nucleation ($n < n_{\rm nuc}$) and $g_n = g_{\rm elong}$ during elongation ($n_{\rm nuc} \le n < N-1$), where $n_{\rm nuc}$ is the critical nucleus size. Finally, inserting the last subunit makes the maximum number of possible contacts and thus enjoys the most favorable association free energy $g_{\rm N}$ (see the Appendix for discussion of an irreversible final assembly step). Except for the last subunit, this model corresponds to Eq. 1.

Our NG model is quite similar to that of Zlotnick and co-workers (2,9), which reproduces experimental assembly kinetics for several viruses (2,5). Nucleation has been shown to correspond to completion of polygons (e.g., a pentamer of dimers for cowpea chlorotic mottle virus (3) or a trimer of dimers for turnip crinkle virus (30)), which can be understood by noting that the subunits assembling to form the first polygon make fewer contacts on average than later subunits. Furthermore, intertwining of flexible terminal arms and other subunit conformation changes can provide additional stabilization upon polygon formation. In Zlotnick and co-workers' formulation (2,9), the nucleation and elongation phases are distinguished by having different forward rate constants, with a slow rate constant before nucleation and a rate constant which is several orders-of-magnitude higher during growth. However, based on the fact that nucleation seems to correspond to the subunit stabilization that results from completion of polygons, we choose to distinguish the nucleation and growth phases with different subunit-subunit association free energies.

We find that our results concerning lag times and identification of the critical nucleus size are insensitive to parameter values. We will therefore present results in Numerical Results for one parameter set: nucleation size $n_{\rm nuc} = 5$ (a pentamer of dimers) and free energy parameters $g_{\rm nuc} = 7\ k_{\rm B}T$, $g_{\rm elong} = 2g_{\rm nuc}$ and $g_{\rm N} = 2g_{\rm elong}$, which imply that adding a subunit becomes, on average, twice as favorable after nucleation and four times as favorable for the final subunit. We note that the value of f corresponds to an average association rate, and not the association rate constant for a single subunit binding site. To clarify this point, we consider a model in the Supporting Material in which we relax the assumption that the association rate constant is independent of intermediate size.

The classical nucleation theory model

We also consider a definition of transition rates based on the classical nucleation theory (henceforth referred to as CNT) suggested by Zandi et al. (31), in which each partial-capsid intermediate is described as a sphere, with a missing spherical cap. The unfavorable free energy due to unsatisfied subunit-subunit interactions at the perimeter of the cap is represented by a line tension σ, and the binding free energy is

$$G_n = ng_{\rm c} + \sigma l_n, \tag{7}$$

with the perimeter of the missing spherical cap given by

$$l_n = 2[\pi n(N-n)/N]^{1/2}, \tag{8}$$

with $g_{\rm c}$ the binding free energy per subunit (not per contact) in a complete capsid. Following Zandi et al. (31), we set the line tension to $\sigma = -g_{\rm c}/2$, which indicates that, on average, a subunit adding to the perimeter of the capsid satisfies half of its contacts. We find that Eq. 7 with $\sigma = -g_{\rm c}/2$ is a reasonable description during the growth phase in BD simulations, but fewer contacts are made by subunits associating to form the first polygon. Note that whereas Eq. 7 is derived under a continuum approximation, we use it to calculate free energies for intermediates with discrete sizes n when solving Eq. 6. We assume that the forward rate constant is proportional to the number of subunits on the perimeter, $f_n = f_0 l_n$, with f_0 the association rate constant for a single binding site. From BD simulation trajectories (14,16,20), we know that this relation overpredicts the net forward rate constant, as only the subunit binding sites that lead to more than one subunit-subunit contact lead to productive assembly for parameters that give rise to effective assembly. However, we show in the Supporting Material that our conclusions are not affected by changing the proportionality constant. As for the NG model, our conclusions about lag times and identifying the critical nucleus are insensitive to parameter values.

The most significant difference between the CNT model and the nucleation and growth models is the size of the critical nucleus. For the NG model the critical nucleus size is $n_{\rm nuc}$, provided $\exp(g_{\rm nuc}/k_{\rm B}T) < c_0 v_0 < \exp(g_{\rm elong}/k_{\rm B}T)$. For the CNT model, the critical nucleus size varies with subunit concentration, and is given by the maximum in $k_{\rm B}T\log(c_1 v_0)n + G_n$, or (31)

$$n_{\mathrm{nuc}} = 0.5N\left(1 - \frac{\Gamma}{\left(\Gamma^2 + 1\right)^{1/2}}\right), \quad (9)$$

with $\Gamma = [g_c - \ln(c_1 v_0)]/\sigma$.

NUMERICAL RESULTS

We explore the concentration dependence of assembly kinetics by numerically integrating the system of rate equations (Eq. 6) with the initial condition $c_1 = c_0$ over a range of subunit concentrations. We start with the NG model, with representative parameters: the contact free energy $g_{\mathrm{nuc}} = 7\ k_{\mathrm{B}}T$ ($\approx$ 4 kcal/mol) (11), capsid size $N = 120$ corresponding to 120 dimer subunits in hepatitis B virus (11), the critical nucleus size $n_{\mathrm{nuc}} = 5$, and the subunit association rate constant $f = 10^5\ \mathrm{M}^{-1}\ \mathrm{s}^{-1}$ (5). As is evident from Eqs. 2 and 4, assembly dynamics are insensitive to g_{elong} and g_{N}, provided that the forward reaction is favorable during elongation; we use $g_{\mathrm{elong}} = 2g_{\mathrm{nuc}}$ and $g_{\mathrm{N}} = 4g_{\mathrm{nuc}}$ (see Rate Equation Models for Capsid Assembly for further discussion).

As shown in Fig. 2, the fraction of subunits in complete capsids ($P_{\mathrm{N}}(t)$) obtained from the rate equations mirrors the results of the BD simulations (Fig. 1 *a*). In both cases, the kinetic trap point c_{kt} roughly corresponds to the point at which the time to build the capsid becomes longer than the average nucleation time (Eq. 5). For the rate equation models, the kinetic trap concentration can be calculated as a function of our parameter values using Eqs. 2, 3, and 5 to give $c_{\mathrm{kt}} \approx 60$, which corresponds well to the point at which assembly times increase with subunit concentration. Assembly eventually occurs above c_{kt} as large partial capsids scavenge subunits from smaller intermediates.

We calculate light-scatter signal I_{LS} from the mass-averaged molecular weight of assemblages (2). Because light-scatter units are arbitrary, we normalize calculated light scatter by N to give 1 if all subunits are in complete capsids ($P_{\mathrm{N}} = 1$). As shown in Fig. 2 *b*, the calculated light scatter closely tracks the completion fraction for initial subunit concentrations below $c_{\mathrm{c}} \approx 38\ \mu\mathrm{M}$, as the majority of assembled subunits are found in complete capsids once the lag phase is complete.

The concentration-dependence of median assembly times and lag times

The median assembly times (reaction half-lives), and lag times numerically calculated for the nucleation and growth model, with respect to the completion fraction P_{N} and calculated light scatter, are given in Fig. 4. As illustrated in Fig. 3, lag times are extracted from numerical solutions for each parameter set with a linear fit to the assembly trace at the point of maximal assembly rate (dP_{N}/dt or dI_{LS}/dt). The lag time is given by the intersection of the linear fit with the baseline value, which is 0 for the completion fraction and roughly $1/N$ for calculated light scatter. Note that for initial subunit concentrations below c_{c}, lag times for both the completion fraction and calculated light scatter are inversely proportional to c_0, and the mean first-passage time estimate (Eq. 2) for the capsid elongation phase τ_{elong} closely predicts the lag time for completion fractions. The lag time for light scatter is shorter than for the completion fraction because signal is integrated over all assemblages, but demonstrates the same scaling. Hence, lag times measured for P_{N} (SEC) or light scattering are associated with the elongation phase, or the time required to build a complete capsid. Note that our expressions for the lag time differ from the well-known expression for the time lag in nucleation theory (32), which gives the time for the concentration of nuclei to reach steady state.

The correspondence between light scatter and completion fraction lag times begins to break down at the crossover concentration c_{c} (Eq. 5), when the concentration of free subunits is depleted before the first capsids finish assembling. Above this point, the molecular weight average growth rate becomes dominated by dimerization and hence varies inversely with the square of initial subunit concentration. The relationship between the lag time and the elongation phase of capsid assembly discovered here explains the observation of Endres and Zlotnick (9) that the duration of lag time is proportional to the elongation forward rate constant.

The reaction half-life (median assembly time) $\tau_{1/2}$ is given by $P_{\mathrm{N}}(\tau_{1/2}) = 0.5P_{\mathrm{N}}^{\mathrm{eq}}$ for the completion fraction, or the analogous relation for calculated light scatter. Below c_{c}, half-lives measured with light scatter and completion fraction agree quantitatively, as anticipated from Fig. 2 *b*, and agree closely with the two-state nucleation kinetics estimate (Eq. 4). In Fig. 4 *a*, the numerical and theoretical half-lives are normalized by the equilibrium $P_{\mathrm{N}}^{\mathrm{eq}}$ to emphasize that the scaling with concentration identifies the critical nucleus size:

$$\tau_{1/2}/P_{\mathrm{N}}^{\mathrm{eq}} \propto c_0^{n_{\mathrm{nuc}}-1}.$$

The fact that assembly times can be predicted from nucleation kinetics alone can be understood by noting that assembly times are dominated by nucleation below c_{c}. The close correspondence between the theoretical and numerical median assembly times below c_{c} suggests that this quantity may provide a simple alternative to the critical nucleus estimator presented in the Appendix of Endres and Zlotnick (9).

As evident from Eq. 4, the critical nucleus size can also be identified by evaluating $\ln\ \tau_{1/2}/P_{\mathrm{N}}$ as a function of the subunit-subunit association free energy g_{nuc}. We find that agreement between the theoretical predictions and numerical results is insensitive to g_{nuc} and n_{nuc}.

Maximum assembly rates

As noted by Endres and Zlotnick (9), it is not reliable to relate the critical nucleus size to initial assembly rates or the extent of assembly as a function of time due to the

presence of the lag phase. At low concentrations, though, the maximum assembly rates approach the initial nucleation rate given in Eq. 3 (see Fig. S10 in the Supporting Material). However, because the maximum assembly rates deviate from the theoretical prediction well below the crossover concentration c_c, it appears that median assembly times are a more robust predictor of the critical nucleus size.

The classical nucleation model

To evaluate whether our conclusions are model-dependent, we also consider the concentration dependence of assembly times and lag times for the CNT model. To facilitate comparisons between the models, we set $g_c = 13.8\ k_BT$ and $f_0 = 10^5/[2(N/\pi)^{1/2}]$, so that P_N^{eq} and the average association rate constants are the same for both models. This choice enables the elongation time prediction of Eq. 2 to be directly compared with the numerically calculated lag time; as shown in Fig. 5, the numerical results and theoretical estimate agree. Furthermore, as for the NG models, the calculated light scatter closely tracks the completion fraction below the kinetic trap point. As noted in Rate Equation Models for Capsid Assembly, the most significant difference between the CNT and NG models is that the critical nucleus size in the classical nucleation model varies with subunit concentration (Eq. 9), with $4 < \hat{n}_{nuc} < 9$ for the simulated range of initial subunit concentrations in this work. The reaction half-life, however, appears to scale roughly as $\tau_{1/2}/P_N \propto C_0^{-9}$ because the effective critical nucleus size increases over the course of the reaction as free subunits are depleted.

For all parameter sets we considered, the CNT model demonstrates a large effective critical nucleus size as the concentration is reduced below c_c. Hence, analyzing the concentration dependence of experimental median assembly times could be one way to evaluate which of the CNT model or NG model better represents capsid assembly mechanisms.

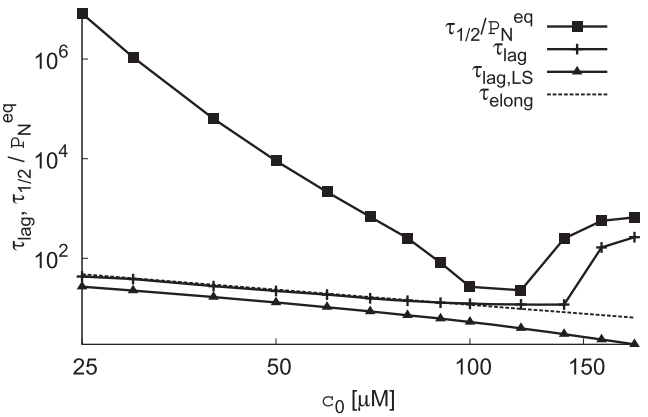

FIGURE 5 The median assembly time (■) and lag time (+) calculated from the completion time and the lag time for light scatter (▲) for the classical nucleation model are shown as functions of concentration. The theoretical prediction for lag time Eq. 2 is shown as a dashed line.

The effect of binding between partial-capsid intermediates

Because binding between intermediates does occur in the BD simulations, their results validate the predictions we have just explored with rate equation models. This fact can be understood by noting that binding between partial capsid intermediates is most prevalent above c_c, when many capsids are assembling at the same time. Although intermediate binding can, in principle, ameliorate the kinetic trapping predicted by the simple models above c_{kt}, in practice the kinetic trap point is only slightly shifted because binding between large partial capsids is rarely compatible with their geometries. Although the time to assemble the first capsid continues to decrease with increasing subunit concentration for the BD simulations (Fig. 1), assembly of subsequent capsids is impeded at the highest subunit volume fractions (Fig. 1 *a*). Furthermore, malformed capsids become increasingly likely as concentrations increase and partial capsids interact (see (16,33) for further discussion of these points).

Binding between intermediates can also arise when the association rates or binding free energies are stronger along one subunit-subunit interface. For most viruses, in fact, the capsid protein rapidly dimerizes and exists primarily as a dimer in solution; the subunit considered in the NG and CNT models corresponds to a protein dimer, and the monomer is not explicitly considered. For systems in which rate constants are arranged so that hierarchical assembly and addition of individual subunits occur with comparable rates, as explored in Sweeney et al. (25), the correspondence between lag times and capsid elongation times will still hold, but the scaling of elongation times (and hence lag times) with subunit concentrations could differ. This possibility would be interesting to examine further.

The slow approach to equilibrium

We note that accurate estimation of P_N^{eq} is not necessary to obtain the critical nucleus size from $\tau_{1/2}$, as nucleation rates have such a dramatic concentration dependence even at subunit concentrations for which $P_N^{eq} > 0.5$. This observation could be important, as the median assembly times shown in Fig. 4 *a* and Fig. 5 extend beyond experimental feasibility at low concentrations (see Fig. S9 in the Supporting Material), particularly for the CNT model because of the large effective critical nucleus size. Furthermore, as noted by Zlotnick (8), capsid assembly approaches equilibrium asymptotically. Therefore, when estimating P_N^{eq}, it is important to plot $P_N(t)$ on a logarithmic scale to judge equilibration, and it may be necessary to extract the contact free energy g_{nuc} or g_c from fits to kinetic data over a series of concentrations (2,5). The slow approach to equilibrium due to nucleation barriers was noted from simulations in Hagan and Chandler (16), and is discussed for the classical nucleation model in Morozov et al. (10).

DISCUSSION AND OUTLOOK

In this work, we examine simulations and two theoretical models for capsid assembly, for which we find that the duration of the lag phase measured by SEC or dynamic light scatter is related to the time for a nucleated capsid to grow to completion, and hence scales inversely with initial subunit concentration. When nucleation of new partial capsids is faster than this growth time, the system becomes kinetically trapped due to starvation of free subunits, meaning that capsid formation rates decreases with increasing initial subunit concentration. If there is a well-defined critical nucleus size, it can be identified from the scaling of the median assembly time with respect to initial subunit concentration.

These predictions have important implications for obtaining mechanistic information about capsid elongation (growth after nucleation) from bulk in vitro assembly kinetics experiments. The fact that overall reaction times are closely predicted by an expression based solely on nucleation kinetics (Eq. 4, Fig. 4) when assembly is most efficient, suggests that the lag phase contains the most information about elongation. In fact, the preceding analysis demonstrates that the average elongation time, and hence the average growth velocity during elongation, are directly related to the duration of the lag phase. Because Chen et al. (6) recently showed that time-resolved light scattering data can be obtained with millisecond resolution, it will be possible to measure lag times with high precision if the reaction is started in a controlled manner. Additional mechanistic information about the elongation process could be obtained if the distribution of growth times could be deconvolved from the distribution of nucleation times. The kinetic trap criterion, however, limits this possibility by constraining growth times, and hence causes the lag phase duration to be short compared to nucleation times.

This constraint could potentially be overcome in several ways. The distribution of elongation times can be directly measured in experiments that monitor the assembly of individual capsids, as the elongation and nucleation phases can be separated (34). For bulk assembly studies, recent theoretical studies (14,35) found that robust assembly is possible under conditions of fast heterogeneous nucleation if there is excess capsid protein. Thus, experimental systems in which capsid assembly is induced by nucleic acids (36), synthetic polymers (37,38), nanoparticles (39), and portal or scaffolding proteins (40–42) could be used to elucidate elongation mechanisms (although assembly mechanisms can be influenced by the heterogeneous component (14,36,43)).

Finally, we note that processes not considered in this work, such as transitions between assembly-active and assembly-inactive conformations of free subunits (6) or hierarchical assembly (44) would add additional complexity to analysis of the lag phase. In particular, if conformational transitions are slow compared to elongation times, the conformational transition timescale will be reflected in the duration of the lag phase. We plan to investigate that case further. In addition, a systematic comparison of model predictions with experimental assembly data over a wide range of concentrations could reveal additional features of complexity in assembly mechanisms and suggest model improvements.

APPENDIX

We consider two additional cases for elongation times in this Appendix.

Elongation times with intermediate-dependent forward rate constants

The parameter f in Eq. 2 is the average forward rate constant for the elongation phase. For the CNT model, the forward rate constants are proportional to the perimeter of the partial capsid intermediate, $f_n = f_0 l_n$, with the perimeter l_n given by Eq. 8. For $fc_1 \gg b_{elong}$ we show in the Supporting Material that $t_{elong} = 0.5(\pi N)^{1/2}/(f_0 c_1)$. In Fig. 5, we see that elongation times for the CNT model can be approximated with Eq. 2 by setting $f = 2f_0(N/\pi)^{1/2}$.

Irreversible elongation steps

It is evident from Eq. 2 that elongation times would be essentially the same if the final subunit insertion step was irreversible. In fact, as shown by Zlotnick (28), all features of assembly dynamics are the same in the case of a final irreversible assembly step except for times greater than the time for subunit dissociation from a complete capsid (b_N). Subunit exchange from complete capsids has been seen experimentally (45), but over a timescale of days rather than the minutes or hours in which assembly takes place. This observation is captured by all of the models considered here, in which insertion of the final subunit is the most favorable free energy, and hence dissociation of a subunit from a complete capsid ($b_N \propto \exp(g_N/k_BT)$) is slow. For the case of several irreversible steps during the elongation phase, elongation times could be calculated by breaking Eq. 2 into independent random walks in series, with one walk between each set of irreversible steps.

SUPPORTING MATERIAL

Five figures are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(09)01752-4.

We are grateful to Jane Kondev for a discussion that led us to investigate the origins of the lag phase and to Robert Jack for discussions about the steady-state ensemble.

M.F.H. and O.M.E. were supported by award No. R01AI080791 from the National Institute of Allergy and Infectious Diseases. M.F.H. also acknowledges support by the National Science Foundation through the Brandeis Materials Research Science and Engineering Center. The BD simulations were performed on the Brandeis High Performance Computing Cluster.

REFERENCES

1. Prevelige, Jr., P. E., D. Thomas, and J. King. 1993. Nucleation and growth phases in the polymerization of coat and scaffolding subunits into icosahedral procapsid shells. *Biophys. J.* 64:824–835.
2. Zlotnick, A., J. M. Johnson, …, D. Endres. 1999. A theoretical model successfully identifies features of hepatitis B virus capsid assembly. *Biochemistry.* 38:14644–14652.

3. Zlotnick, A., R. Aldrich, …, M. J. Young. 2000. Mechanism of capsid assembly for an icosahedral plant virus. *Virology*. 277:450–456.
4. Casini, G. L., D. Graham, …, D. T. Wu. 2004. In vitro papillomavirus capsid assembly analyzed by light scattering. *Virology*. 325:320–327.
5. Johnson, J. M., J. H. Tang, …, A. Zlotnick. 2005. Regulating self-assembly of spherical oligomers. *Nano Lett*. 5:765–770.
6. Chen, C., C. C. Kao, and B. Dragnea. 2008. Self-assembly of brome mosaic virus capsids: insights from shorter time-scale experiments. *J. Phys. Chem. A*. 112:9405–9412.
7. Berthet-Colominas, C., M. Cuillel, …, B. Jacrot. 1987. Kinetic-study of the self-assembly of brome mosaic-virus capsid. *Eur. Biophys. J. Biophys. Lett*. 15:159–168.
8. Zlotnick, A. 1994. To build a virus capsid. An equilibrium model of the self-assembly of polyhedral protein complexes. *J. Mol. Biol*. 241:59–67.
9. Endres, D., and A. Zlotnick. 2002. Model-based analysis of assembly kinetics for virus capsids or other spherical polymers. *Biophys. J*. 83:1217–1230.
10. Morozov, A. Y., R. F. Bruinsma, and J. Rudnick. 2009. Assembly of viruses and the pseudo law of mass action. *J. Chem. Phys*. 131:155101.
11. Ceres, P., and A. Zlotnick. 2002. Weak protein-protein interactions are sufficient to drive assembly of hepatitis B virus capsids. *Biochemistry*. 41:11525–11531.
12. Kegel, W. K., and P. Schoot Pv. 2004. Competing hydrophobic and screened-Coulomb interactions in hepatitis B virus capsid assembly. *Biophys. J*. 86:3905–3913.
13. Bar-Haim, A., and J. Klafter. 1998. On mean residence and first passage times in finite one-dimensional systems. *J. Chem. Phys*. 109:5187–5193.
14. Hagan, M. F. 2008. Controlling viral capsid assembly with templating. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys*. 77:051904.
15. Schwartz, R., P. W. Shor, …, B. Berger. 1998. Local rules simulation of the kinetics of virus capsid self-assembly. *Biophys. J*. 75:2626–2636.
16. Hagan, M. F., and D. Chandler. 2006. Dynamic pathways for viral capsid assembly. *Biophys. J*. 91:42–54.
17. Nguyen, H. D., V. S. Reddy, and C. L. Brooks, 3rd. 2007. Deciphering the kinetic mechanism of spontaneous self-assembly of icosahedral capsids. *Nano Lett*. 7:338–344.
18. Wilber, A. W., J. P. K. Doye, …, P. Wong. 2007. Reversible self-assembly of patchy particles into monodisperse icosahedral clusters. *J. Chem. Phys*. 127:085106.
19. Hicks, S. D., and C. L. Henley. 2006. Irreversible growth model for virus capsid assembly. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys*. 74:031912.
20. Elrad, O. M., and M. F. Hagan. 2008. Mechanisms of size control and polymorphism in viral capsid assembly. *Nano Lett*. 8:3850–3857.
21. Rapaport, D. C. 2008. Role of reversibility in viral capsid growth: a paradigm for self-assembly. *Phys. Rev. Lett*. 101:186101.
22. Nguyen, H. D., V. S. Reddy, and C. L. Brooks. 2009. Invariant polymorphism in virus capsid assembly. *J. Am. Chem. Soc*. 131:2606–2614.
23. Endres, D., M. Miyahara, …, A. Zlotnick. 2005. A reaction landscape identifies the intermediates critical for self-assembly of virus capsids and other polyhedral structures. *Protein Sci*. 14:1518–1525.
24. Zhang, T. Q., and R. Schwartz. 2006. Simulation study of the contribution of oligomer/oligomer binding to capsid assembly kinetics. *Biophys. J*. 90:57–64.
25. Sweeney, B., T. Zhang, and R. Schwartz. 2008. Exploring the parameter space of complex self-assembly through virus capsid models. *Biophys. J*. 94:772–783.
26. Maibaum, L. 2008. Phase transformation near the classical limit of stability. *Phys. Rev. Lett*. 101:256102.
27. Ceres, P., S. J. Stray, and A. Zlotnick. 2004. Hepatitis B virus capsid assembly is enhanced by naturally occurring mutation F97L. *J. Virol*. 78:9538–9543.
28. Zlotnick, A. 2007. Distinguishing reversible from irreversible virus capsid assembly. *J. Mol. Biol*. 366:14–18.
29. Reference deleted in proof.
30. Sorger, P. K., P. G. Stockley, and S. C. Harrison. 1986. Structure and assembly of turnip crinkle virus. II. Mechanism of reassembly in vitro. *J. Mol. Biol*. 191:639–658.
31. Zandi, R., P. van der Schoot, …, H. Reiss. 2006. Classical nucleation theory of virus capsids. *Biophys. J*. 90:1939–1948.
32. Wu, D. T. 1992. The time-lag in nucleation theory. *J. Chem. Phys*. 97:2644–2650.
33. Whitelam, S., E. H. Feng, …, P. L. Geissler. 2009. The role of collective motion in examples of coarsening and self-assembly. *Soft Matter*. 5:1251–1262.
34. Jouvenet, N., P. D. Bieniasz, and S. M. Simon. 2008. Imaging the biogenesis of individual HIV-1 virions in live cells. *Nature*. 454:236–240.
35. Hagan, M. F. 2009. A theory for viral capsid assembly around electrostatic cores. *J. Chem. Phys*. 130:114902.
36. Johnson, J. M., D. A. Willits, …, A. Zlotnick. 2004. Interaction with capsid protein alters RNA structure and the pathway for in vitro assembly of cowpea chlorotic mottle virus. *J. Mol. Biol*. 335:455–464.
37. Sikkema, F. D., M. Comellas-Aragonès, …, R. J. Nolte. 2007. Monodisperse polymer-virus hybrid nanoparticles. *Org. Biomol. Chem*. 5:54–57.
38. Hu, Y., R. Zandi, …, W. M. Gelbart. 2008. Packaging of a polymer by a viral capsid: the interplay between polymer length and capsid size. *Biophys. J*. 94:1428–1436.
39. Sun, J., C. DuFort, …, B. Dragnea. 2007. Core-controlled polymorphism in virus-like particles. *Proc. Natl. Acad. Sci. USA*. 104:1354–1359.
40. Parent, K. N., S. M. Doyle, …, C. M. Teschke. 2005. Electrostatic interactions govern both nucleation and elongation during phage P22 procapsid assembly. *Virology*. 340:33–45.
41. Parent, K. N., A. Zlotnick, and C. M. Teschke. 2006. Quantitative analysis of multi-component spherical virus assembly: scaffolding protein contributes to the global stability of phage P22 procapsids. *J. Mol. Biol*. 359:1097–1106.
42. Tuma, R., H. Tsuruta, …, P. E. Prevelige. 2008. Detection of intermediates and kinetic control during assembly of bacteriophage P22 procapsid. *J. Mol. Biol*. 381:1395–1406.
43. McPherson, A. 2005. Micelle formation and crystallization as paradigms for virus assembly. *Bioessays*. 27:447–458.
44. Misra, N., D. Lees, …, R. Schwartz. 2008. Pathway complexity of model virus capsid assembly systems. *Comput. Math. Method Med*. 9:277–293.
45. Parent, K. N., M. M. Suhanovsky, and C. M. Teschke. 2007. Phage P22 procapsids equilibrate with free coat protein subunits. *J. Mol. Biol*. 365:513–522.