

Relaxational dynamics of the T -number conversion of virus capsids

Cite as: J. Chem. Phys. 159, 084904 (2023); doi: 10.1063/5.0160822

Submitted: 5 June 2023 • Accepted: 7 August 2023 •

Published Online: 23 August 2023



Alexander Bryan Clark,¹  Mohammadamin Safdari,¹  Selim Zoorob,²  Roya Zandi,^{1,2,a)} 
and Paul van der Schoot³ 

AFFILIATIONS

¹ Department of Physics and Astronomy, University of California, Riverside, California 92521, USA

² Biophysics Graduate Program, University of California, Riverside, California 92521, USA

³ Department of Applied Physics and Science Education, Eindhoven University of Technology, Postbus 513, 5600 MB Eindhoven, The Netherlands

^{a)} Author to whom correspondence should be addressed: royaz@ucr.edu

ABSTRACT

We extend a recently proposed kinetic theory of virus capsid assembly based on Model A kinetics and study the dynamics of the interconversion of virus capsids of different sizes triggered by a quench, that is, by sudden changes in the solution conditions. The work is inspired by *in vitro* experiments on functionalized coat proteins of the plant virus cowpea chlorotic mottle virus, which undergo a reversible transition between two different shell sizes ($T = 1$ and $T = 3$) upon changing the acidity and salinity of the solution. We find that the relaxation dynamics are governed by two time scales that, in almost all cases, can be identified as two distinct processes. Initially, the monomers and one of the two types of capsids respond to the quench. Subsequently, the monomer concentration remains essentially constant, and the conversion between the two capsid species completes. In the intermediate stages, a long-lived metastable steady state may present itself, where the thermodynamically less stable species predominate. We conclude that a Model A based relaxational model can reasonably describe the early and intermediate stages of the conversion experiments. However, it fails to provide a good representation of the time evolution of the state of assembly of the coat proteins in the very late stages of equilibration when one of the two species disappears from the solution. It appears that explicitly incorporating the nucleation barriers to assembly and disassembly is crucial for an accurate description of the experimental findings, at least under conditions where these barriers are sufficiently large.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0160822>

I. INTRODUCTION

There are two major aspects to how viruses in general, and more specifically, unenveloped single-stranded RNA (ssRNA) viruses, infect living cells. The first is the role of the viral genome, which is responsible for the replication of the virus inside the host cell. The second is that of the capsid, the protein shell that protects the genome of the virus until it enters the cell in order to release its genome and produce large numbers of virus particles from the replicated genome and protein components. Therefore, virus capsids must be able to function as a protective transporter as well as a deliverer of genomes. These two functions must be executed via controlled shell disassembly and reassembly, depending on the environment and stage in the reproduction cycle.^{1–5}

Experimentally, in order to study the physical aspects of the disassembly and reassembly of viruses, use is often made of a simple,

icosahedral plant virus known as Cowpea Chlorotic Mottle Virus (CCMV). *In vitro* studies have shown that CCMV disassembles spontaneously under appropriate solution conditions. This occurs when the buffer solution containing the virus particles is suddenly changed from neutral to basic pH at high ionic strength.^{6–8} Conversely, solutions containing the coat proteins of CCMV or those of other viruses such as Brome Mosaic Virus (BMV) and Hepatitis B Virus (HBV) can be made to spontaneously encapsulate negatively charged cargo, which does not need to be the native genome but may include heterologous ssRNAs, synthetic polyanions, or even nanocolloids.^{9–12}

The self-assembled particles often have the same size or structure as the native virus, but this is not always the case.^{13–17} In some cases, mixtures of differently sized or structured particles self-assemble, as is the case for HBV even *in vivo*.^{18,19} *In vitro*, such size competition is often observed too, where the predominance of a

certain particle size, or so-called T -number, appears to depend on the size of the cargo and even on the solution conditions.^{20–22} We note that under appropriate physico-chemical conditions or by modifying the coat proteins, e.g., by removing or modifying the RNA binding domain, empty or filled capsids of different sizes and shapes may also be produced.^{23–27}

The spontaneous conversion between assembled and disassembled states or that between different sizes of assembled particles, with or without cargo, is often modeled in terms of a process akin to a thermodynamic phase transition.^{7,21,22,28–32} This is possible because (i) the number of proteins required to form a shell is sufficiently large, typically many tens to hundreds of proteins, and (ii) the number of intermediates consisting of incomplete assemblies is exceedingly small in comparison to the number of free coat proteins and complete particles.^{33,34} The latter is believed to be caused by a rim tension associated with proteins at the edge that miss contact with neighboring proteins in the incomplete shell.³⁵ Scattering experiments³⁶ and computer simulations seem to support this view.^{2,37}

While we have a reasonable understanding of the thermodynamics of viruses and virus-like particles formed by empty capsids and capsids containing heterologous RNAs, synthetic polyanions, nanocolloids, etc., our understanding of the kinetics of assembly and disassembly of such products remains very sketchy indeed.^{3,38–40} Experimentally, this is in part due to limitations in the spatial and time resolution of the intermediate structures between fully formed capsids and the free subunits in the solution.^{1,36} The situation is even more complicated in the presence of a polyelectrolyte cargo,⁴¹ where different assembly pathways have been identified that are referred to as “concerted” or “en mass” vs “sequential,” which may depend on the solution conditions.^{9,38,42–44} Both seem to be governed by nucleation and growth stages.

Clearly, the assembly and disassembly of empty capsids is, in principle, conceptually the most simple, as it involves only the coat proteins. Even if we focus on the self-assembly kinetics of empty capsids, until recently the experimental focus was mainly on assembly, not disassembly. This is probably due to disassembly being a significantly more sluggish process than assembly, which expresses itself in a strong hysteresis between the assembly and disassembly of virus capsids.^{35,45,46} Available experimental data on a variant of the HBV coat protein suggests that the disassembly pathway from a fully formed capsid to free monomers includes a stage where fractal aggregates are formed, although this may well depend on the solution conditions.⁴⁵ CCMV capsids seem to involve two types of intermediate states, one large and one small.⁴⁰ Again, a lack of spatial and temporal resolution experimentally hampers obtaining a detailed picture of how precisely a virus capsid disintegrates. Computer simulations suggest that, despite its stochastic character, the onset of disassembly may involve only a limited number of distinct contacts between coat proteins.⁴⁷

In a recent study, Timmermans *et al.* functionalized the coat protein of CCMV and studied the reversible conversion between a $T = 3$ capsid at low pH and high ionic strength and a $T = 1$ capsid at high pH and lower ionic strength.⁴⁸ For non-functionalized coat proteins of CCMV, empty $T = 3$ capsids form due to hydrophobic interactions overwhelming the electrostatics in a solution with 500 mM NaCl at a pH of 5.0, while no structures form when the pH is increased and ionic strength decreases.⁴⁹ By functionalizing the

coat protein with elastin-like polypeptides, the hydrophobic interactions of the polypeptides at a higher pH of 7.5 with a lower ionic strength of 100 mM NaCl dominate the electrostatic interactions and form the smaller species of capsid.⁴⁸ This conversion requires the disassembly of one species in order to assemble the other, as most of the coat proteins present in the solution turn out to reside in capsids and not in free proteins. The conversion of the modified CCMV coat proteins from $T = 1$ to $T = 3$ occurs much more slowly than that from $T = 3$ to $T = 1$. To rationalize this finding, Timmermans *et al.* extended a classical nucleation theory for virus capsid assembly of a single capsid size to two capsid sizes, introducing also a disassembly pathway,³⁵ which allowed them to interpret their results in terms of free energy barriers between the free monomers and the $T = 1$ and $T = 3$ assembled states, as well as the differences in the binding free energy gains of assembly.

While classical nucleation theory seems to be able to describe the final assembly products and match the experimental data quite well up to a few days after a quench, that is, a sudden change in solution conditions, the curve fitting procedure turns out to be numerically demanding due to a large number of adjustable system parameters and the scatter in the data. Hence, we ask the question if the experiments of Timmermans *et al.* can also be described by a much simpler model based on what in the theoretical phase transition community is called Model A kinetics.^{50,51} This kind of kinetics is purely relaxational and hinges on the concept of generalized forces in a free energy landscape. It was applied before to describe the assembly and disassembly kinetics of a single type of empty capsid.⁵¹ Here, we extend it to mixtures of monomers and two species of capsid and focus on the conversion between these two types of capsid following a quench. For deep quenches, nucleation barriers are small, and such an approach should be appropriate.

The advantage of Model A kinetic theory is that it involves the deterministic time evolution of the conversions between individual coat protein subunits and fully formed assemblies which can be solved analytically for shallow quenches. Analytical solutions are also possible for deep quenches at experimentally relevant time scales, namely for early and late times. This circumstance allows us to identify driving forces at different stages of the size conversion. Indeed, our calculations show that in the early stages of the conversion, the translational entropy of the free subunits drives the assembly of one of the capsid species, while in the later stages, the impact of differences in the binding free energy of the two species of capsid predominates.

The remainder of this paper is structured as follows: We first reiterate the equilibrium theory for capsid competition and analyze in detail the equation of state describing the amount of proteins free in $T = 1$ and $T = 3$ capsids as a function of the overall concentration of coat proteins and two binding free energies. Next, we write the kinetic equations for the fraction of proteins in the various species within the framework of Model A kinetic theory⁵⁰ and express these in terms of the equilibrium quantities. We next investigate analytically the evolution of the distribution of proteins for shallow quenches when initial and final solution conditions are close together, producing two elementary time scales. By numerically solving the equations, we confirm that under most quench conditions, even if the initial and final states of aggregation are very different, the approach to equilibrium involves two time scales that we are able to interpret. Next, we apply the theory to the experimental

data of Timmermans *et al.*,⁴⁸ where we conclude that for those experiments, Model A kinetics describes the early and intermediate stages of the kinetics reasonably well but fails to reproduce the ratio of assembly products at the final stage of the experiments. It seems that an explicit description of nucleation processes is essential to explaining the experimental findings. In the final section, we summarize our results and discuss under what experimental conditions we expect our theory to work at all times.

II. EQUILIBRIUM THEORY FOR CAPSID COMPETITION

Our first goal is to set up a statistical thermodynamic model describing the competition between fully formed shells of different T -numbers as a function of the solution conditions, which acts to drive the dynamics of the process. Based on the equilibrium theory of capsid assembly, we need only consider the coat protein subunits free in solution and those present in the complete capsids; incomplete intermediate structures (partially formed capsids) are statistically highly improbable on account of a line tension associated with missing neighbors of coat proteins making up the cap rim.^{33–35,51,52}

Presuming the solution is dilute, we write the dimensionless free energy f per coat protein as follows:

$$f = \eta_s \ln c \eta_s - \eta_s + \sum_T \left[\frac{\eta_T}{q_T} \ln \frac{c \eta_T}{q_T} - \frac{\eta_T}{q_T} + \eta_T g_T \right] \quad (1)$$

in terms of the fraction of proteins in free solution η_s and those in the capsid species η_T with $T = 1, 3, 4, 7, \dots$ the triangulation number of the capsids and associated aggregation number q_T . Below, we focus on the case where we only have two species in competition with each other, namely $T = 1$ and $T = 3$, which are the most prevalent. (Competition between pseudo- $T = 2$ and $T = 3$ has been observed in the context of the encapsulation of polyanions by virus coat proteins.⁵³) Further, $c \ll 1$ is the overall mole fraction of coat proteins in solution, and g_T is the mean dimensionless binding free energy of a single coat protein in a complete capsid of size T . For stable capsids to form, the latter must be negative and is typically in the range of -10 to -20 (in units of thermal energy).^{22,54}

The free energy equation (1) is the sum of an ideal mixing entropy (stemming from the translational entropy of each species) and the net binding free energy accounting for the subunit-subunit interactions in a fully assembled capsid.⁵⁴ The optimal distribution of the coat proteins over the various states of assembly minimizes the free energy, requiring that

$$\left(\frac{\partial f}{\partial \eta_T} \right)_{c, g_T, \eta_T = \eta_{T,\infty}} = 0, \quad (2)$$

where we define $\eta_{T,\infty}$ as the value of η_T under conditions of thermodynamic equilibrium. Note that we eliminate the fraction of coat proteins in solution from the free energy by making use of the conservation of mass by inserting the identity $\eta_s = 1 - \sum_T \eta_T$. Demanding Eq. (2) to hold, the equations of state become

$$-\ln \left[c \left(1 - \sum_T \eta_{T,\infty} \right) \right] + \frac{1}{q_T} \ln \frac{c \eta_{T,\infty}}{q_T} + g_T = 0, \quad (3)$$

representing the law of mass action for all potentially present capsid species.

For the case in which we have a competition between two distinct capsids with triangulation numbers T and $T' > T$, mirroring the experiments of Timmermans *et al.*,⁴⁸ Eq. (3) tells us that we need to solve two equations of state self-consistently, namely

$$\ln \left[c \left(1 - \eta_{T,\infty} - \eta_{T',\infty} \right) \right] = \frac{1}{q_T} \ln \frac{c \eta_{T,\infty}}{q_T} + g_T, \quad (4)$$

$$= \frac{1}{q_{T'}} \ln \frac{c \eta_{T',\infty}}{q_{T'}} + g_{T'}. \quad (5)$$

Note that $q_{T'}/q_T = T'/T$, which is larger than unity if, without loss of generality, we presume that $T' > T$. Furthermore, $q_T = 60T$ if the coat proteins are monomers and $q_T = 30T$ if they are dimers, as is the case for CCMV.

In the limit of very small degrees of assembly where most proteins remain in free solution and $\eta_{T,\infty} \ll 1$, Eqs. (4) and (5) can be solved to give rise to the solutions $\eta_T \sim q_T c^{-1} (c/c_{*,T})^{q_T}$ and $\eta_{T'} \sim q_{T'} c^{-1} (c/c_{*,T'})^{q_{T'}}$ with $c_{*,T} = \exp(g_T)$ and $c_{*,T'} = \exp(g_{T'})$ critical concentrations associated with the binding free energies g_T and $g_{T'}$. Since typically $q_T, q_{T'} \gg 1$, the transition between the capsid-poor and capsid-rich states is very sharp indeed. In fact, the larger the T number, the sharper the transition.

Subtracting Eq. (4) from (5) and defining the difference in dimensionless binding free energies of the two species as $\Delta g \equiv g_{T'} - g_T$, we obtain a simple relation between the fraction of protein present in the two species of capsid,

$$\eta_{T,\infty} = \alpha (\eta_{T',\infty})^{T/T'}, \quad (6)$$

where

$$\alpha \equiv \left(\frac{T}{T'} \right)^{T/T'} \left(\frac{q_T}{c} \right)^{1-T/T'} \exp(q_T \Delta g). \quad (7)$$

We immediately see that if $\alpha \ll 1$, which in essence implies $\Delta g < 0$ as $q_T \gg 1$, we must have $\eta_{T,\infty} \ll 1$. In that case, the T' species predominates. The two species are equally prevalent if $\eta_T = \eta_{T'} = \alpha^{T'/(T'-T)}$ provided $\alpha < 2^{-(T'-T)/T'}$ as by definition $\eta_T + \eta_{T'} < 1$. Figure 1 illustrates the predicted competition between $T = 1$ and $T' = 3$ capsids. As shown in the figure, for $\alpha = 2^{-2/3} \simeq 0.63$ at relatively low degrees of assembly, most of the proteins in capsids reside in the $T = 1$ structure. However, if almost all of the proteins present in the solution are capsids, they divide equally among both structures for $\alpha \simeq 0.63$. Since $\alpha \propto c^{-1+T/T'} = c^{-2/3}$, we conclude that, keeping everything else constant but increasing the total protein concentration, the fraction of protein in $T = 1$ capsids decreases relative to that in $T' = 3$ capsids. This conclusion extends to any mixture of capsids with $T' > T$.

Since we have $\eta_{T,\infty}$ as a function of $\eta_{T',\infty}$, we can insert it in Eq. (5) to obtain

$$\ln \left[1 - \alpha (\eta_{T',\infty})^{T/T'} - \eta_{T',\infty} \right] = \frac{1}{q_{T'}} \ln \left[\frac{c \eta_{T',\infty}}{q_{T'}} \right] + \ln \left[\frac{c_{*,T'}}{c} \right], \quad (8)$$

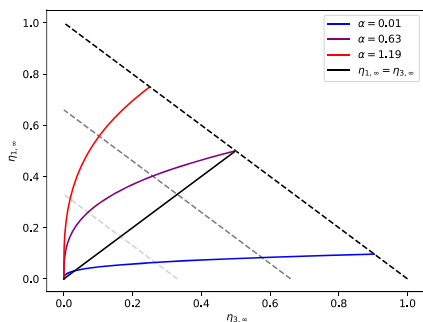


FIG. 1. Relation between the fraction of protein in $T = 1$ capsids, $\eta_{1,\infty}$, and that in $T = 3$ capsids, $\eta_{3,\infty}$ (6), for different values of the parameter α , according to Eq. (7). From top to bottom: $\alpha = 0.01$ (red), $\alpha = 0.63$ (purple), and $\alpha = 1.19$ (blue). The dashed lines demarcate the maximum value of $\eta_{3,\infty}$ for every value of $\eta_{1,\infty}$ for different maximum fractions of subunits present in capsids: $\eta_1 + \eta_3 = 1$, $2/3$, and $1/3$ from top to bottom. The solid black line describes the special case $\eta_{3,\infty} = \eta_{1,\infty}$.

for which we have not been able to find an exact analytical solution. However, if $\eta_{T',\infty}$ is not exceedingly small, the first term on the right-hand side of the equation is much smaller than the second, and we can write

$$1 - \alpha(\eta_{T',\infty})^{T'/T} - \eta_{T',\infty} = \left(\frac{c_{*,T'}}{c} \right). \quad (9)$$

This equation can be solved analytically for some values of the ratio T'/T . For instance, for the competition between $T = 1$ and $T = 3$ capsids, we can write Eq. (9) as a cubic equation for the quantity $\eta_{3,\infty}^3$. Instead of presenting lengthy expressions, we here focus on the limiting behavior of Eq. (9) to extract useful information.

Equation (9) reveals that if $\eta_{3,\infty} \gg \alpha^{3/2}$, we have $\eta_{3,\infty} \sim 1 - (c_{*,3}/c)$ provided that $c \gtrsim c_{*,3}$, and if $c \lesssim c_{*,3}$, we must have $\eta_{3,\infty} \approx 0$. Interestingly, for the case in which $\eta_{3,\infty} \ll \alpha^{2/3}$, we find that $\eta_{3,\infty}$ once more becomes very small. In that case, we again retrieve Eq. (9) from Eq. (8), except that $c_{*,3}$ will be replaced by $c_{*,1}$, leading to the asymptotic relationship $\eta_{1,\infty} \sim 1 - (c_{*,1}/c)$ for $c \gtrsim c_{*,1}$, while for $c \lesssim c_{*,1}$ we obtain $\eta_{1,\infty} \approx 0$.

In conclusion, if $c_{*,3} \ll c_{*,1}$ and $c \lesssim c_{*,3}$, almost no capsids form. However, for $c \gtrsim c_{*,3}$ capsids do form, but they are mostly $T = 3$ structures, even if $c \gtrsim c_{*,1}$. If the coat protein concentration is sufficiently large and $c_{*,1} \approx c_{*,3}$, both species form in appreciable quantities. Notice that for $c_{*,1} \approx c_{*,3}$ to hold, the difference in binding free energies of the two species must be much smaller than unity, $|\Delta g| = |g_3 - g_1| \ll 1$. As the binding free energies tend to be in the range of 10 to 20 in units of thermal energy, very small differences in binding free energy are required to see the co-existence between different T numbers under conditions of thermodynamic equilibrium.

Figures 2 and 3 illustrate the competition between the two capsid sizes as a function of the binding free energies g_1 , g_3 , and Δg at a coat protein concentration (mole fraction) of $c = 2 \times 10^{-5}$. The coat proteins are presumed to be dimeric, mimicking the experiments of Timmermans *et al.* performed with the modified CCMV coat proteins.⁴⁸ To this end, we set $q_1 = 30$ and $q_3 = 90$. For the given concentration of coat protein, we expect capsids to disappear from the

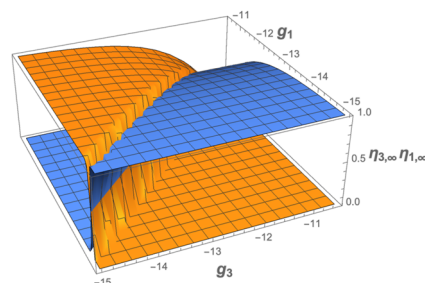


FIG. 2. Equilibrium fraction of proteins in $T = 1$ capsids, $\eta_{1,\infty}$, and in $T = 3$ capsids, $\eta_{3,\infty}$, as a function of the binding free energies of the two structures, g_1 and g_3 [see Eqs. (6) and (8)]. The orange surface shows $\eta_{3,\infty}$, and the blue surface shows $\eta_{1,\infty}$. The dimensionless overall protein concentration was set at $c = 2 \times 10^{-5}$, and the aggregation numbers for the two species were set to values of $q_1 = 30$ and $q_3 = 90$, implying that the protein subunits are dimers of the coat protein. See also the main text.

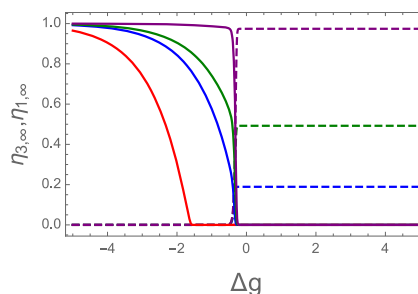


FIG. 3. Equilibrium fraction of protein in $T = 3$ capsids, $\eta_{3,\infty}$ (solid lines), and $T = 1$ capsids, $\eta_{1,\infty}$ (dashed lines), as a function of the difference in the dimensionless binding free energies $\Delta g \equiv g_3 - g_1$. Shown are results for different values of g_1 , indicated by different colors, with purple for $g_1 = -14$, green for $g_1 = -11$, blue for $g_1 = -10.4$, and red for $g_1 = -9$. The dimensionless overall protein concentration was set at $c = 2 \times 10^{-5}$, and the aggregation numbers were set equal to $q_1 = 30$ and $q_3 = 90$, implying that the subunits are dimers of the coat proteins. Note that the transition does not occur at $\Delta g = 0$. See also the main text.

solution if g_1 and g_3 are greater than $-\ln 2 \times 10^{-5} \approx -11$, as is confirmed in Fig. 2. According to Eq. (6), the crossover from the $T = 1$ to $T = 3$ dominated regimes occurs for $\eta_{1,\infty} = \eta_{3,\infty} = 1/2 = \alpha^{3/2}$. This translates to a difference between the binding free energies of $\Delta g = g_3 - g_1 \approx -0.3$. The lack of crossover for $\Delta g = 0$ is due to the impact of entropy, which favors smaller capsids. Figure 3 confirms this fact and shows $\eta_{1,\infty}$ and $\eta_{3,\infty}$ as a function of Δg for different values of $g_1 = -14, -11, -10.4, -9$. We expect little assembly for $g_1 \gtrsim -10$ because the concentration then drops below the corresponding critical concentration. The figure confirms this trend for both T numbers.

III. KINETICS OF T -NUMBER CONVERSION

Having obtained a clear understanding of the thermodynamics of mixtures of differently sized capsids, we now have a theory that aims to describe the time evolution of the solution composition following a quench. Since we are interested in the deterministic time evolution of the conversion between capsid species, we write

the corresponding kinetic equations in terms of what is known as Model A relaxational kinetics, following our earlier work on single capsid species assembly and disassembly.⁵¹ The governing equations are for the most general case, involving multiple species of capsids, given by

$$\frac{\partial \eta_T}{\partial t} = -\Gamma_T \frac{\partial f}{\partial \eta_T}, \quad (10)$$

where Γ_T are phenomenological rate constants associated with the various T -numbers. As the quench experiments are done at fixed concentrations,⁴⁸ we need not consider their dependence on the total concentration of coat proteins c . The possibility that the rate constants depend on the concentration of free monomers cannot be excluded, in particular the one for the species that assembles. We choose to ignore this and view the rate constants as adjustable parameters, as is usually done in this kind of dynamical model. Notice that because of the mass conservation $\eta_s = 1 - \sum_T \eta_T$, the time evolution of the free monomers in solutions $\partial \eta_s / \partial t = -\sum_T \partial \eta_T / \partial t$ can be easily obtained.

Investigating the conversion between two capsid species, we focus specifically on the conversion between $T = 1$ and $T = 3$ capsids. By inserting our free energy, Eq. (1), into the kinetic equation (10) for $T = 1$ and $T = 3$, and making use of the equations of state for the two species of capsid, Eqs. (4) and (5), we obtain

$$\frac{\partial \eta_1}{\partial \tau} = \ln \left(\frac{1 - \eta_1 - \eta_3}{1 - \eta_{1,\infty} - \eta_{3,\infty}} \right) - \frac{1}{q_1} \ln \frac{\eta_1}{\eta_{1,\infty}}, \quad (11)$$

$$\frac{\partial \eta_3}{\partial \tau} = \Gamma \left[\ln \left(\frac{1 - \eta_1 - \eta_3}{1 - \eta_{1,\infty} - \eta_{3,\infty}} \right) - \frac{1}{q_3} \ln \frac{\eta_3}{\eta_{3,\infty}} \right]. \quad (12)$$

Here, we have made the time t dimensionless by defining $\tau \equiv \Gamma_1 t$ and introduced the kinetic parameter $\Gamma \equiv \Gamma_3 / \Gamma_1$ that depends on the relative rates of assembly and disassembly of the $T = 3$ and $T = 1$ capsids. As before, $\eta_{T,\infty}$ denotes the fraction of coat proteins in each species under conditions of thermodynamic equilibrium after the quench, as time goes to $\tau \rightarrow \infty$ (hence the subscript ∞).

While we have not been able to solve these coupled differential equations exactly, approximate analytical solutions can be found within certain experimentally relevant limits. For instance, for shallow quenches or the late stages of the conversion for deep quenches, we can set $\eta_T(\tau) = \eta_{T,\infty}(1 + \delta_T(\tau))$ where $|\delta_T(\tau)| \ll 1$. Inserting this into Eqs. (11) and (12), Taylor expands these in terms of $|\delta_T(\tau)|$ giving rise to a set of linear equations that can be diagonalized. We refer to Appendix A for details. The two fundamental relaxation rates $\lambda_{\pm}$ obtained through solving the coupled equations are complicated functions of the aggregation numbers and the equilibrium values of the fraction of proteins in the two species. We here only quote their approximate values,

$$\lambda_+ \sim \frac{1 + \Gamma}{\eta_{s,\infty}} + \dots, \quad (13)$$

where $\eta_{s,\infty} = 1 - \eta_{1,\infty} - \eta_{3,\infty}$ is the fraction of coat proteins in free solution, and

$$\lambda_- \sim \frac{\Gamma}{1 + \Gamma} \left[\frac{1}{\eta_{1,\infty} q_1} + \frac{1}{\eta_{3,\infty} q_3} \right] + \dots \quad (14)$$

to leading order in powers of $1/\eta_{1,\infty} q_1 \ll 1$ and $1/\eta_{3,\infty} q_3 \ll 1$. Notice that in the limits $\Gamma \rightarrow 0$ and $\Gamma \rightarrow \infty$, one relaxation becomes infinitely faster than the other, as one would expect. Figure 4, showing the numerically obtained values for the case $q_3 = 3q_1 \gg 1$, confirms this.

It appears that one of the relaxation rates, λ_+ , depends only on the final total amount of assembled material, or, equivalently, on the fraction of free monomers, not on how the material is distributed over the two competing capsid species. This suggests that this rate describes the response of the free monomers. The other rate, $\lambda_- \leq \lambda_+$, does depend on the final distribution over the two capsid species and is not symmetrical with respect to their fractions. It is tempting to assign this rate to the late-stage equilibration involving mainly the two types of capsid and less so that of the free protein subunits.

We believe that this explains most of our numerical results discussed in more detail in Sec. IV, namely that the fraction of proteins in monomers relaxes relatively fast to its final value, after which the assembly of one type of capsid becomes enslaved by the disassembly of the other. Obviously, this would suggest that in that case, we can ignore the first term on the right-hand side of Eqs. (11) and (12), making the kinetic equations amenable to an analytical solution.⁵¹ However, these solutions cannot be correct because then η_1 and η_3 would evolve independently from each other. Indeed, these solutions violate the equality $\partial \eta_1 / \partial \tau = -\partial \eta_3 / \partial \tau$ that holds if the concentration of free subunits $\eta_s = 1 - \eta_1 - \eta_3$ was constant.

For initial conditions where $\eta_T(0)/\eta_{T,\infty}$ are not very large or very small and for sufficiently short times $\tau \ll \tau_- \equiv \lambda_-^{-1}$, we are however able to find an analytical solution for the fraction of proteins in free solution $\eta_s = 1 - \eta_1 - \eta_3$. For this, we first ignore the last terms in Eqs. (11) and (12) and obtain $\partial \eta_1 / \partial \tau = \Gamma^{-1} \partial \eta_3 / \partial \tau = \ln(\eta_s(\tau)/\eta_{s,\infty})$. Subsequently, combining these two equations gives a dynamical equation for η_s that can be solved exactly to give

$$\left(\frac{1 + \Gamma}{\eta_{s,\infty}} \right) \tau = \text{li} \left[\frac{\eta_s(0)}{\eta_{s,\infty}} \right] - \text{li} \left[\frac{\eta_s(\tau)}{\eta_{s,\infty}} \right], \quad (15)$$

where $\text{li}[x] = \int_0^x dy (1/\ln y)$ is the logarithmic integral. For the first equality, we used the identity given in Eq. (13). This confirms that

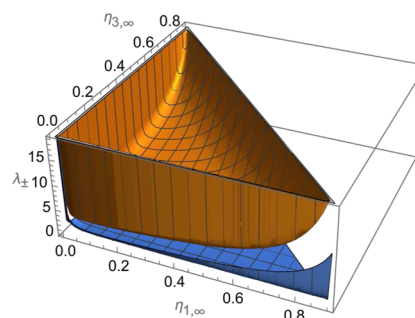


FIG. 4. Dimensionless fundamental relaxation rates $\lambda_{\pm}$ of the linearized kinetic equations (11) and (12) describe the conversion between $T = 1$ and $T = 3$ capsids as a function of the fraction of $T = 3$ capsids η_3 and kinetic coefficient $\Gamma = 1$. The blue surface shows λ_- , and the orange one λ_+ . We set $q_3 = 3q_1 = 90$ to model CCMV capsid coat proteins that are present in the solution as dimers.

λ_+ must indeed be the relaxation rate associated with the early-stage kinetics dominated by the free monomers and one of the capsid species, in particular if $\Gamma \gg 1$ or $\Gamma \ll 1$.

Consequently, the initial response to the quench must be driven primarily by the translational entropy of the free subunits and the binding free energy of the capsid species. In contrast, the late-stage relaxation, in which there are protein exchanges between the two types of capsids, is basically driven by differences in the binding free energies of the proteins in these two species. Our numerical results, discussed in Sec. IV, confirm this picture, albeit with notable exceptions.

IV. NUMERICAL RESULTS

We solved the coupled differential equations (11) and (12) numerically using the SciPy library `scipy.integrate`,⁵⁵ investigating the interconversion of $T = 1$ and $T = 3$ capsids. We systematically varied all input parameters, that is, (i) the starting compositions $\eta_1(0)$ and $\eta_3(0)$, (ii) the final compositions $\eta_{1,\infty}$ and $\eta_{3,\infty}$, and (iii) the ratio of assembly rates Γ , taking as representative values 0.1, 1, and 10. To select the appropriate time interval $d\tau$ for our numerical studies, we demand that the concentration changes after each time step be less than 1%. If the concentration change exceeds 1%, we decrease the time interval. Our results turn out to be invariant of the time interval step for $d\tau < 10^{-3}$ for all cases investigated. For our discussion below, we set its value at $d\tau = 10^{-4}$.

We focus in this section on three interesting types of numerical experiments: (i) one starting with a large fraction of proteins equally distributed in both capsids and ending with a large fraction of proteins only in one of the capsids (see Fig. 5); (ii) one where the initial state is one with almost no capsids present in solution but where most proteins are in capsids in the final (equilibrium) state (Fig. 6); and (iii) one starting with a large fraction of proteins only in one of the capsids and ending with a large fraction of proteins in the other capsid (Fig. 7). We discuss these three general classes using representative examples.

Figure 5 shows the time evolution of the various species present in the solution from the initial conditions $\eta_1(0) = \eta_3(0) = 0.40$ to the final states $\eta_{1,\infty} = 0.80$ and $\eta_{3,\infty} = 0.18$. So, we start off with equal amounts of 40% coat proteins in $T = 1$ and $T = 3$ capsids and 20% in free solution, ending up with much more protein in $T = 1$ than in $T = 3$, that is, 80% vs 18%, and almost no free protein. Independent of the rate Γ that defines how swiftly the material in $T = 3$ capsids responds to changes in the thermodynamic conditions relative to that in $T = 1$ capsids, the fraction of free monomers indicated by the green curves in Fig. 5 decreases monotonically to reach its final steady-state value well before the proteins in capsids have reached a steady state at longer times. We observed this behavior in over 100 cases that we investigated. This also seems to be consistent with the outcome of the experiments of Timmermans *et al.*, who also observed a nearly constant monomer fraction from the earliest time they were able to do the measurements.⁴⁸ See also Sec. V.

If $\Gamma = 0.1$, the response of the fraction of proteins in $T = 3$ capsids is relatively slow, while for $\Gamma = 10$, it is fast. This means that if $\Gamma = 0.1$, the $T = 1$ capsids, which are thermodynamically more stable than the $T = 3$ capsids, quickly absorb monomers from the solution to increase their number, whereas the $T = 3$ capsids do not disassemble yet on account of their slow dynamics. This is what the dotted

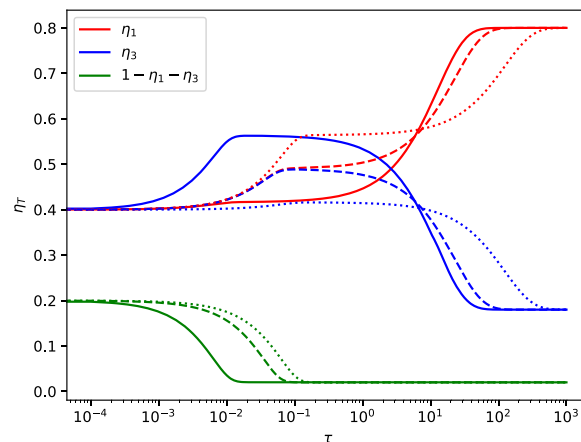


FIG. 5. Time evolution of the fraction of coat proteins η_T in capsids with triangulation numbers $T = 1$ (red) and $T = 3$ (blue), as well as the fraction in monomers $1 - \eta_1 - \eta_3$ (green), as a function of dimensionless time τ . Indicated are results for different values of the ratio of assembly rates: $\Gamma = 10$ (solid), $\Gamma = 1$ (dashed), and $\Gamma = 0.1$ (dotted). The initial conditions are $\eta_1(0) = \eta_3(0) = 0.40$, and the final states are $\eta_{1,\infty} = 0.80$ and $\eta_{3,\infty} = 0.18$. We set $q_3 = 3q_1 = 90$, assuming the coat proteins form dimers. See also the main text.

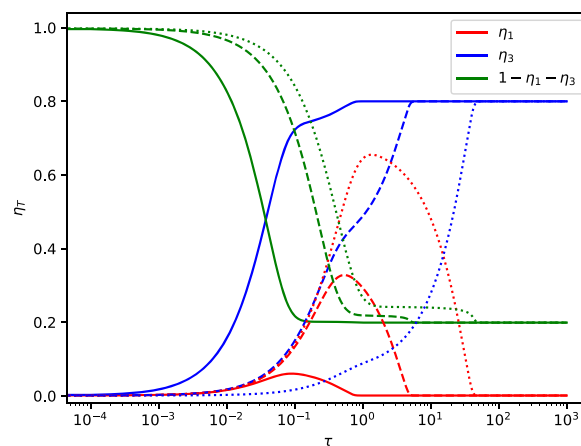


FIG. 6. Time evolution of the fraction of coat proteins η_T in capsids with triangulation numbers $T = 1$ (red) and $T = 3$ (blue), as well as the fraction in monomers $1 - \eta_1 - \eta_3$ (green), as a function of dimensionless time τ . Indicated are results for different values of the ratio of assembly rates: $\Gamma = 10$ (solid), $\Gamma = 1$ (dashed), and $\Gamma = 0.1$ (dotted). The initial conditions are $\eta_1(0) = \eta_3(0) = 0.001$, and the final states are $\eta_{1,\infty} = 0.001$ and $\eta_{3,\infty} = 0.80$. We set $q_3 = 3q_1 = 90$, assuming the coat proteins form dimers. See also the main text.

red and blue curves in Fig. 5 indicate. The figure also reveals that at a much later time, after η_1 has reached a pseudo-plateau, $T = 3$ shells start to disassemble in order to form additional $T = 1$ shells at a more or less constant fraction of free coat proteins in solution. So, for this small value of Γ , the conversion of $T = 3$ into $T = 1$ capsids happens in two steps, first involving the formation of $T = 1$ capsids using free monomers, and subsequently $T = 1$ capsids form upon the disassembly of $T = 3$ capsids.

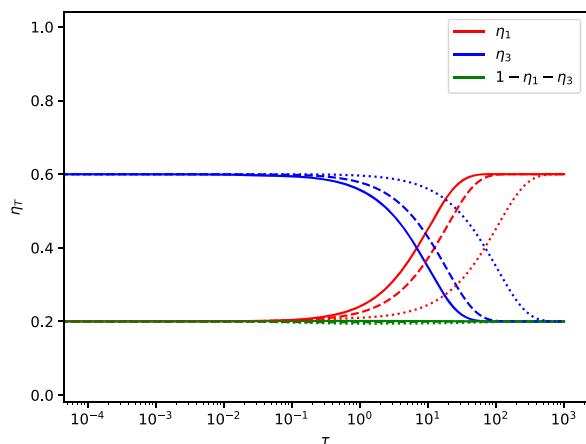


FIG. 7. Time evolution of the fraction of coat proteins η_T in capsids with triangulation numbers $T = 1$ (red) and $T = 3$ (blue), as well as the fraction in monomers $1 - \eta_1 - \eta_3$ (green), as a function of dimensionless time τ . Indicated are results for different values of the ratio of assembly rates: $\Gamma = 10$ (solid), $\Gamma = 1$ (dashed), and $\Gamma = 0.1$ (dotted). The initial conditions are $\eta_1(0) = 0.2$ and $\eta_3(0) = 0.6$, and the final states are $\eta_{1,\infty} = 0.6$ and $\eta_{3,\infty} = 0.2$. We set $q_3 = 3q_1 = 90$, assuming the coat proteins form dimers. See also the main text.

Something similar happens for $\Gamma = 10$, shown in Fig. 5 by the solid curves, but now reversed. In this case, even though the $T = 3$ shell is less stable than the $T = 1$ shell, both capsids are more stable than free monomers. Thus, in the beginning, $T = 3$ shells start to assemble, thereby depleting the free monomers in the solution. Next, η_3 reaches a pseudo-plateau, where it remains constant for quite some time, after which the $T = 3$ shells disassemble in favor of the $T = 1$ shells, which in the end are more thermodynamically stable. This happens at a more or less constant free monomer concentration. In this case, the fraction of $T = 3$ capsids first increases from its initial value and then decreases to a value lower than the initial value. Our results for $\Gamma = 1$ are similar to those for $\Gamma = 10$, except that the overshoot of η_3 is much smaller and happens much later.

A careful analysis in Fig. 5 thus reveals that there must be two time scales involved, not just in the linear response discussed in Sec. III; see Eqs. (13) and (14), but also in the full, non-linear response of the system of coat proteins and capsid shells to a quench. If we translate the relaxation rates $\lambda_{\pm}$ of Eqs. (13) and (14), obtained from a linear response analysis, to relaxation times $\tau_{\pm} \equiv \lambda_{\pm}^{-1}$, then we find $\tau_+ \simeq 0.018$ and $\tau_- \simeq 110$ for $\Gamma = 0.1$, and $\tau_+ \simeq 0.0018$ and $\tau_- \simeq 11$ for $\Gamma = 10$. This roughly matches the time scales of the two processes, as may be verified in Fig. 5.

The picture that emerges remains valid even if we start off with mostly free monomers, as illustrated in Fig. 6. Here, we monitor the time evolution of the various species present in the solution from the initial conditions $\eta_1(0) = \eta_3(0) = 0.001$ to the final states $\eta_{1,\infty} = 0.001$ and $\eta_{3,\infty} = 0.80$. So, we initiate the assembly with equal amounts of 0.1% coat protein in $T = 1$ and $T = 3$ capsids and 99.8% in free solution, ending in much more protein in $T = 3$ than in $T = 1$ with 80% vs 0.1% and 19.9% free protein.

We note that if $\Gamma = 0.1$, the response of the fraction of proteins in $T = 3$ shells is relatively slow, while for $\Gamma = 10$, it is fast. Hence, if $\Gamma = 0.1$, $T = 1$ capsids swiftly assemble by absorbing free monomers

from the solution, even though the $T = 3$ shells are the more stable of the two. The thermodynamic stability of the $T = 1$ capsids exceeds that of the free monomers, driving the assembly of the metastable species. As the dotted line in Fig. 6 shows, after some time, these particles disassemble again in favor of the $T = 3$ shells. This happens at a more or less constant concentration of free monomers. Again, we see that the competition in time involves only two species: first, the proteins in $T = 1$ capsids and free solution respond to the quench, and after that, the protein subunits in $T = 1$ and $T = 3$ shells readjust to reach a state of thermodynamic equilibrium.

It transpires that for the given conditions, $T = 1$ shells appear only temporarily. The larger the value of Γ , the less pronounced this effect is. The appearance and disappearance of the smaller species produce a shoulder and a pseudo-plateau in the fraction of protein in the larger species, not dissimilar to what we saw in Fig. 5. Again, two time scales appear that we associate with the elementary rates $\lambda_{\pm}$. Translated to relaxation times, we obtain $\tau_+ \simeq 0.026$ and $\tau_- \simeq 12$ for $\Gamma = 0.1$ and $\tau_+ \simeq 0.012$ and $\tau_- \simeq 0.25$ for $\Gamma = 10$. This roughly matches the time scales of the two processes, as revealed in Fig. 6. Note that to obtain these numbers, we used the full expression given in Eq. (A2) in Appendix A rather than the asymptotic Eqs. (13) and (14) as $\eta_{1,\infty} q_1$ is not large enough for these asymptotic expressions to hold.

The final case that we discuss is that where $\eta_1(0) = 0.2$ and $\eta_3(0) = 0.6$, and $\eta_{1,\infty} = 0.6$ and $\eta_{3,\infty} = 0.2$. So, we are deeply in the polymerized regime, with 80% of monomers in capsids: 60% in $T = 3$ shells and 20% in $T = 1$ shells at time zero. In thermal equilibrium, these numbers are reversed, but their total fraction remains the same. This turns out to be a special case. Figure 7 reveals that, even though the starting and ending fractions of free monomers are the same, the conversion of $T = 3$ into $T = 1$ capsids occurs at an *approximately* constant free monomer fraction. For $\Gamma = 10$, there is a slight overshoot (solid line), while for $\Gamma = 1$ and $\Gamma = 0.1$, there is a slight undershoot (dashed and dotted lines). The over- and undershoots are very small indeed, at most a few percent.

Again, there are two time scales, one of which signifies the onset of the very small overshoot or undershoot of the free monomers. This happens well before the fraction of protein in capsids responds significantly. The second (much larger) time scale is associated with the conversion between the large and smaller species of capsid and with the relaxation of the monomer concentration back to its equilibrium value. The corresponding relaxation times we obtain from the rates of Eqs. (13) and (14) are $\tau_+ \simeq 0.18$ and $\tau_- \simeq 100$ for $\Gamma = 0.1$ and $\tau_+ \simeq 0.018$ and $\tau_- \simeq 10$ for $\Gamma = 10$. These numbers agree again approximately with what is shown in Figs. 7 and 8, where we zoom in on the time evolution of the fraction of free monomers. The latter figure also points to the existence of relatively long-lived pseudo-plateaus characterized by out-of-equilibrium concentrations of free monomers.

The question arises: what happens if we swap the initial and final fractions for the case shown in Fig. 7. We expect from the fundamental relaxation times $\tau_{\pm}$ that this situation will be quite different: τ_+ does not change since the fraction of free protein does not change, but the second time scale changes considerably to $\tau_- = 60$ for $\Gamma = 0.1$ and $\tau_- = 6.0$ for $\Gamma = 10$ compared to $\tau_- \simeq 100$ and $\tau_- \simeq 10$, respectively. So, the dynamics do change and cannot be compensated for by simply taking the reciprocal value of Γ . See also Fig. 8, showing the time evolution of the fraction of monomer units

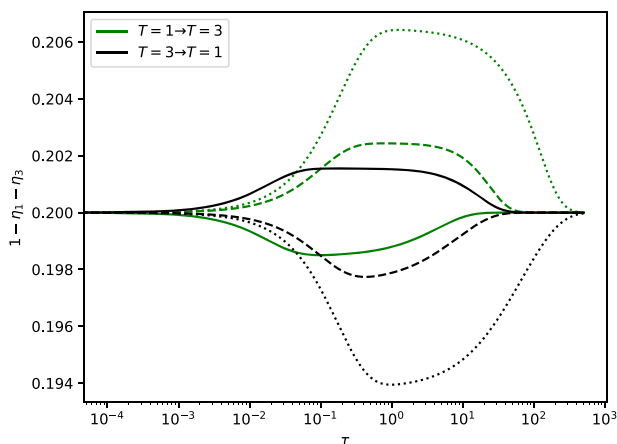


FIG. 8. Time evolution of the fraction of monomers $1 - \eta_1 - \eta_3$, as a function of dimensionless time τ . Indicated are results for different values of the ratio of assembly rates: $\Gamma = 10$ (solid), $\Gamma = 1$ (dashed), and $\Gamma = 0.1$ (dotted). The initial conditions for $T = 3 \rightarrow T = 1$ (black) are $\eta_1(0) = 0.2$ and $\eta_3(0) = 0.6$, with final states $\eta_{1,\infty} = 0.6$ and $\eta_{3,\infty} = 0.2$. For $T = 1 \rightarrow T = 3$ (green), we swap the initial and final conditions of $T = 3 \rightarrow T = 1$. We set $q_3 = 3q_1 = 90$, assuming the coat proteins form dimers. See also the main text.

in free solution for the $T = 1 \rightarrow T = 3$ (green lines) and $T = 3 \rightarrow T = 1$ (black lines) conversions. This makes the assembly and disassembly dynamics fundamentally asymmetric, which is also the case for a single species.⁵¹ The effect is not very pronounced, although, as already mentioned and as a comparison of the different curves in Fig. 8 shows. The most remarkable difference between the cases $T = 1 \rightarrow T = 3$ (green lines) and $T = 3 \rightarrow T = 1$ (black lines) is in the over- and undershoots, which are reversed for corresponding values of Γ .

This ends our discussion of the numerical evaluation of the dynamical equations. We next apply the theory in order to describe the experiments of Timmermans *et al.*⁴⁸

V. COMPARISON WITH EXPERIMENT

In the experiments of Timmermans *et al.*, aqueous solutions containing CCMV coat proteins functionalized with a hydrophobic elastin-like polypeptide form $T = 3$ capsids at a pH equal to 5 and an ionic strength of 0.5M, whereas at a pH of 7.5 and an ionic strength of 0.1M, they assemble into the smaller $T = 1$ capsids.⁴⁸ The sizes were established using a combination of size exclusion chromatography (SEC) and transmission electron microscopy. Changing (by means of a dialysis step) the acidity and ionic strength from pH 7.5 and 0.1M to pH 5 and 0.5M leads to the slow conversion of smaller capsids into larger ones, as is evidenced by SEC. Conversely, SEC also shows that changing the acidity from pH 5 to pH 7.5 and the ionic strength from 0.5 to 0.1M leads to the conversion of the larger species into the smaller species. Since partial capsids are neither observed in the SEC traces nor in the electron micrographs, the conclusion is that the unstable species disassembles into protein subunits (dimers) that reassemble into stable species.

Both types of conversions are exceedingly slow to complete. Starting off at 100% $T = 3$ capsids, after 168 h, only 70% of $T = 1$

capsids are produced, indicating the conversion has not yet been completed at that time. One final data point at 1608 h shows 92% $T = 1$ capsids. The initial response to the change in solution conditions is relatively swift; however, after about 2 h, the rate of conversion slows down considerably. Something rather similar happens in the $T = 1 \rightarrow T = 3$ conversion experiments: starting off at 100% $T = 1$, the process slows down after about 24 h. After 168 h, the conversion is only 60% complete. The final data point at 1440 h shows 98% conversion. In the absence of data between 168 and the final measurements at 1608 and 1440 h for the two types of experiments, we cannot be certain how long the conversion between species actually takes, even if we treat the last points (both taken after about two months) as essentially complete.

While our model does produce a short and a long time scale, mimicking what Timmermans *et al.* observe experimentally, we have not been able to accurately describe the conversion for times below 168 h or to get a complete or near-complete conversion after two months. In this respect, our model is lacking when compared to classical nucleation theory,⁴⁸ which nicely describes the experiments. However, if we take the data up to 168 h at face value, assuming that the system has reached the steady state, and ignore the full conversion that happens between 168 h and two months, we can describe the experimental data rather well.

To fit our model to the experimental data of Timmermans *et al.*,⁴⁸ we minimize the sum of the cost functions, or equivalently, maximize the coefficient of determination r^2 , between the numerical solutions to our coupled equations and the experimental observations for each species. We define the cost function as $\Sigma(\eta_{\text{theo}} - \eta_{\text{exp}})^2 / \Sigma(\eta_{\text{exp}} - \bar{\eta}_{\text{exp}})^2 = 1 - r^2$, where the sum is over all data points for each species, η_{theo} are the fitted values obtained from the theory, η_{exp} are the experimental data, and $\bar{\eta}_{\text{exp}} = \Sigma \eta_{\text{exp}} / N$ where the sum is again over all data points for each species, with $N = 20$ being the number of experimental data points for each conversion experiment. Note that we calculate r_1^2 for the conversion of $T = 1$ to $T = 3$ and r_3^2 for the conversion of $T = 3$ to $T = 1$. We then maximize the average value of $(r_1^2 + r_3^2)/2$. The optimization is accomplished by implementing a hyperparameter optimization method for the grid search variant.⁵⁶

The values for Γ_1 and Γ follow from the hyperparameter optimization and the subsequent model fitting. In order to do the model fitting, we took the initial conditions to be the average of the experimental measurements for the $T = 1$ and $T = 3$ species. We ran a grid search for the equilibrium conditions and their corresponding rate constants.⁵⁷ The search for $\eta_{1,\infty}$ and $\eta_{3,\infty}$ spanned from the smallest to the largest species fraction measurements. More specifically, we set one limit of the search to be $\eta_T(0)$ and the other limit to be $\eta_T(30 \text{ h})$. We then partitioned this so that the step is 0.01. The value of $\eta_{T,\infty}$ was chosen in the range of experimental data. The rate constants Γ_1 and Γ were determined by an extensive search of possible values. When determining the cost function, we first scaled the time axis such that $t\Gamma_1 = \tau$ and then found Γ in order to numerically solve Eqs. (11) and (12). To convert back to experimental time, we take $t = \tau/\Gamma_1$.

In our curve-fitting procedure, we take the initial and final fractions to match the values experimentally observed in the time frame from 0 to 168 h and optimize the phenomenological rate constant ratio Γ . It is important to point out that the actual zero time in the experiments is not known accurately on account of the experimental

procedure that involves a dialysis step. We refer to the original publication of Timmermans *et al.* for experimental details.⁴⁸ The SEC data are consistent with the presence of a fraction of small protein subunits, identified as coat protein dimers, two larger species of particles associated with the $T = 1$ and $T = 3$ capsids, and much larger particles that could be aggregates of (incomplete) capsids or protein subunits. The fraction of protein aggregates remains practically constant over the course of time, as does the fraction of free subunits. We count the material in aggregates as part of the fraction of subunits. We do not expect this to significantly affect the number of capsids in view of the long time scales involved in assembly and disassembly. Hence, we take the SEC's data at face value.

Figure 9 shows the results of the combined sets of experiments for the conversion of $T = 1$ to $T = 3$ particles. In addition, indicated are our model fits (solid lines) to the indicated experimental data (symbols). To obtain the curve fits, we set the initial fractions at $\eta_1(0) = 0.86$ and $\eta_3(0) = 0.02$, and the equilibrium fractions at $\eta_{1,\infty} = 0.42$ and $\eta_{3,\infty} = 0.44$. From the latter and using Eq. (3), we conclude that the binding free energies must be equal to $g_1 \approx -11.4$ and are $g_3 \approx -11.7$, indicating $\Delta g = -0.3$. The curve fitting produces an optimal ratio of rate constants $\Gamma = \Gamma_3/\Gamma_1 = 23.5$, where the corresponding rate constant for the $T = 1$ species Γ_1 is 1.0 h^{-1} , using a time step of $dt = 0.0001 \text{ h}$. The agreement between theory and experiment is fair, with an r^2 value of 0.884. We observe that the agreement is less accurate in the later stages compared to the initial stages. The theory appears to overlook the slight increase in free protein subunit concentration that occurs in the later stages of the process. On the other hand, the scatter in the data is quite large, and the two experimental datasets shown only agree approximately with each other.

Figure 10 shows the results of the quench experiments, which began with a solution primarily containing $T = 3$ particles and progressed to conditions where the most dominant species are the $T = 1$

ones. Again, the symbols indicate the two sets of experimental data, and the solid lines indicate the fits of the data. We set the initial fractions at $\eta_1(0) = 0.01$ and $\eta_3(0) = 0.72$, and the equilibrium fractions at $\eta_{1,\infty} = 0.50$ and $\eta_{3,\infty} = 0.30$. From the latter, we obtain binding energies $g_1 \approx -11.1$ and $g_3 \approx -11.3$, indicating that $\Delta g = -0.2$. We extract $\Gamma = 0.60$ and $\Gamma_1 = 15.2 \text{ h}^{-1}$, using a time step $d\tau = 0.0001 \text{ h}$. The agreement between theory and experiment has significantly improved, with an r^2 value of 0.909, even though the scatter in the data remains quite substantial. Notice that we do recover a virtually constant fraction of protein subunits. Table I shows the rate constants and binding free energies obtained from our data fitting analysis of the interconversion of $T = 1$ and $T = 3$.

If we compare the binding free energies g_1 and g_3 that we obtain with those from the nucleation theory in Ref. 48, we find that our values are somewhat smaller in magnitude. The slightly less negative values that we find should not come as a complete surprise, as our analysis aimed to produce a pseudo-plateau for times approaching 168 h, which we considered to be equilibrium. In contrast, Timmermans *et al.* considered the fractions of the two capsid species after two months as equilibrium values, by which time the less stable species had largely disappeared from the solution.

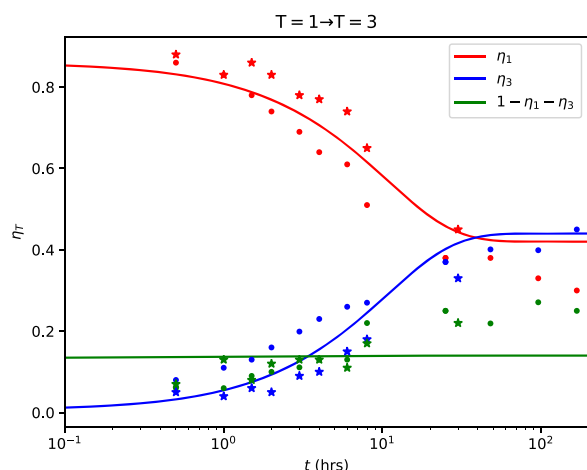


FIG. 9. Fraction of protein in various species η_T as a function of time t in hours. Symbols: results from two datasets where, following a quench, $T = 3$ capsids convert into $T = 1$ capsids. Only the first 168 h of the measurements by Timmermans *et al.*⁴⁸ are shown. The red, blue, and green lines indicate the fractions of $T = 1$, $T = 3$, and free subunits. See also the main text.

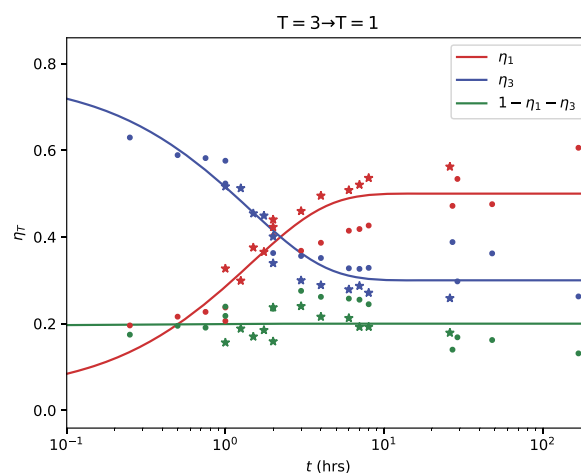


FIG. 10. Fraction of protein in various species η_T as a function of time t in hours. Symbols: results from two datasets where, following a quench, $T = 1$ capsids convert into $T = 3$ capsids. Only the first 168 h of the measurements by Timmermans *et al.*⁴⁸ are shown. The red, blue, and green lines indicate the fractions of $T = 1$, $T = 3$, and free subunits. See also the main text.

TABLE I. Numerical values of the model parameters obtained from fitting the relaxation model to the experimental data for the conversions of $T = 1 \rightarrow T = 3$ and $T = 3 \rightarrow T = 1$ capsids with $d\tau = 0.01$. Γ_1 and Γ_3 are the relaxation rates associated with the assembly and disassembly of the $T = 1$ and $T = 3$ capsids, and $\Gamma = \Gamma_3/\Gamma_1$ is their ratio. The dimensionless free energies of binding of coat protein subunits are g_1 and g_3 , and $\Delta g = g_3 - g_1$ is their difference. See also the main text.

Conversion	$\Gamma_1 \text{ (h}^{-1}\text{)}$	$\Gamma_3 \text{ (h}^{-1}\text{)}$	Γ	g_1	g_3	Δg
$T = 1 \rightarrow T = 3$	1.0	23.5	23.5	-11.4	-11.7	-0.3
$T = 3 \rightarrow T = 1$	15.2	9.1	0.60	-11.1	-11.3	-0.2

We furthermore conclude that for the $T = 1 \rightarrow 3$ conversion experiments, $\Gamma_3 > \Gamma_1$, whereas for the $T = 3 \rightarrow 1$ conversion, $\Gamma_3 < \Gamma_1$. So, for the former type of experiment, the relaxation rate associated with the $T = 3$ particle is larger than that associated with the $T = 1$ particle, and for the latter, the reverse is the case. Since Γ_T is a phenomenological rate that somehow incorporates forward and backward rates, this finding is difficult to interpret. The fundamental relaxation time associated with the free monomers that can be deduced from Eq. (13) is just below 0.01 h (so under one minute) for both types of experiments, which explains why the prediction for the fraction of free monomer subunits remains constant on the time scale of the experiments.

Having completed our discussion of the comparison between theory and experiment, we next summarize our findings and discuss in greater detail the conditions under which our model is likely to be accurate or inaccurate.

VI. DISCUSSION AND CONCLUSIONS

In this work, we propose a purely relaxational, Model A type kinetic theory in order to describe how, as a function of time, a sudden change in the solution conditions modifies the preferred capsid size in aqueous dispersions containing virus coat proteins that assemble into capsids of varying sizes under different physico-chemical conditions. Our aim is to shed light on the underlying competition between protein subunits and the two types of capsids, one of which has become thermodynamically unstable following the quench, while the other has become stable but is not yet present in appreciable quantities in the solution. Our calculation was inspired by the results of time-resolved experiments by Timmermans *et al.*, who studied the quench-induced interconversion of $T = 1$ and $T = 3$ capsids formed by coat proteins of CCMV functionalized by attaching elastin-like polypeptides.⁴⁸

According to our theory, there are essentially two relevant time scales. In almost all of the situations we investigated, there was a time scale associated with the response of the coat protein subunits in free solution to the assembly or disassembly of one of the two types of capsids. The other, typically much longer time scale, describes the equilibration of the two types of capsid under conditions of almost constant free subunit concentration. This implies that, in that case, protein subunits shed by the unstable species are used up by the production of the other species. Consequently, the assembly of the more stable species becomes enslaved by the disassembly of the less stable species. Comparison with the experimental findings of Ref. 48 supports the existence of this scenario. We note that the classical nucleation theory applied to competing capsids also points to a swift equilibration of the monomer concentration.⁴⁸

The theory, although simple, predicts the existence of long-lived metastable states in which the thermodynamically less stable capsid species can become more prevalent over time and even predominate over the more stable species. This is reminiscent of the nucleation of crystals, where metastable crystal phases may nucleate before stable ones do, often explained in terms of the Ostwald “rule of stages,” which states that the phase that nucleates is (in some sense) the one closest in free energy to the parent phase.⁵⁸ The term “closest” may also be interpreted in the kinetic sense, that is, in terms of what phase is the most rapidly accessible, which for crystals is thought to be determined by a trade-off between the strength

of the thermodynamic driving force and the height of the nucleation barrier separating the unstable phase and the more stable solid phase.⁵⁹

In the context of our Model A kinetic theory of the size conversion of polymorphic virus capsids, quasi-stationary states occur when the free subunits have reached their final concentration but the concentrations of the two capsid species have not yet equilibrated. The latter are kinetically controlled not by nucleation barriers, which are absent in the model, but by the ratio of two fundamental relaxation rates. For coat proteins that assemble into $T = 1$ and $T = 3$ capsids, these are the rates Γ_1 and Γ_3 . The relaxation rates Γ_1 and Γ_3 are phenomenological parameters that somehow capture aspects of the underlying molecular processes involved in the assembly and disassembly of both capsid species.

The rate at which the concentration of the thermodynamically most stable species of capsid can respond following a quench is dictated by the relative magnitude of the two relaxation rates, which is the parameter $\Gamma = \Gamma_3/\Gamma_1$ in the model. If the value of Γ is large, the fraction of $T = 3$ capsids may initially grow even if this species is less stable than the $T = 1$ species. And, vice versa, if Γ is small, $T = 1$ capsids may preferentially form even if they are less stable than $T = 3$ capsids. The transient states that ensue may be very long-lived before the thermodynamically more stable species, in the end, take over.

Our theory is able to describe the first 168 h of the experiments of Timmermans *et al.* reasonably well, allowing us to fix all the model parameters. These include the parameters Γ_1 and Γ_3 but also the binding free energies g_1 and g_3 , presented in Table I. For the latter, we find values that are somewhat smaller than those obtained from classical nucleation theory.⁴⁸ The slightly smaller values that we find here are not surprising and result from our curve-fitting procedure, in which we consider the steady state concentrations of species after 168 h as the infinite time concentrations (equilibrium) when both capsid species remain present in appreciable fractions. However, after two months, near-complete conversion of the dominant species occurs in the experiments. If we use as input parameters the fractions of $T = 1$ and $T = 3$ capsids after two months, we are not able to reproduce the experimental data in the time domain from zero to 168 h, as shown in Appendix B.

Superficially, this may seem somewhat surprising, as our model does predict the existence of pseudo-plateaus under appropriate conditions. We have to realize, however, that in order to obtain a near-complete conversion for a long time, it would require an additional time scale. Indeed, from the experiments, we have to conclude that there must be at least three time scales: one involving the protein subunits, describing the time required for their concentration to relax; one involving the redistribution of protein among capsids, leading to a quasi steady state and associated pseudo-plateau; and one leading to the equilibrium state, which essentially eliminates the thermodynamically less stable capsid species from solution. Arguably, the required additional time scale or time scales can be provided by nucleation processes not part of our Model A kinetics, explaining why classical nucleation theory is indeed able to provide a description for the entire capsid conversion process even though it is numerically demanding. We surmise that in the classical nucleation theory put forward in Ref. 48, the additional long time scale is produced by introducing time-dependent, instantaneous concentrations of the various species within a quasi stationary approximation.^{22,60}

The nucleation barriers for the $T = 1 \rightarrow T = 3$ and $T = 3 \rightarrow T = 1$ conversions in solutions of functionalized CCMV coat proteins, required for the assembly of the more stable species and disassembly of the less stable species, are several tens of times greater than the thermal energy.⁴⁸ Clearly, since our model lacks any explicit nucleation barriers, it should not come as a surprise that it cannot describe the full time-dependence of the experiments. Despite this, the model does provide a reasonably accurate description of the available experimental data up to 168 h and turns out to produce good estimates for the binding free energies.

Under conditions where the nucleation barriers are sufficiently small and assembly and disassembly are no longer nucleated, we expect the model to be able to describe the full time dependence. Experimentally, this would require a larger separation in the binding energies of the two capsid species. Practically, this could be achieved, e.g., by varying the properties of the elastin-like polypeptides used in the functionalization of the CCMV coat proteins.⁶¹ Another method to experimentally reduce any nucleation barriers would be to introduce packaging signals, allowing stable structures to form when they would normally be unstable.⁶²

ACKNOWLEDGMENTS

We thank Suzanne Timmermans for the discussions and sharing of the experimental data. We are also grateful to Jordan Teeter for the discussions. A.B.C., M.S., S.Z., and R.Z. acknowledge support from NSF Grant No. DMR-2131963 and the University of California Multicampus Research Programs and Initiatives (Grant No. M21PR3267). P.v.d.S. thanks R.Z. and UC Riverside for hosting his sabbatical and for their hospitality and Eindhoven University of Technology for the financial support.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Alexander Bryan Clark: Data curation (lead); Formal analysis (lead); Methodology (lead); Writing – original draft (equal); Writing – review & editing (equal). **Mohammadamin Safdari:** Data curation (lead); Formal analysis (lead); Methodology (lead); Supervision (equal); Writing – original draft (equal); Writing – review & editing (equal). **Selim Zoorob:** Data curation (equal); Formal analysis (supporting); Methodology (supporting). **Roya Zandi:** Conceptualization (lead); Formal analysis (supporting); Funding acquisition (lead); Project administration (lead); Supervision (lead); Writing – original draft (equal); Writing – review & editing (lead). **Paul van der Schoot:** Conceptualization (lead); Formal analysis (supporting); Supervision (lead); Writing – original draft (lead); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

APPENDIX A: LINEAR ANALYSIS

To investigate the dynamical equations (11) and (12) at the level of a linear response theory, we insert $\eta_1(\tau) = \eta_{1,\infty}(1 + \delta_1(\tau))$ and $\eta_3(\tau) = \eta_{3,\infty}(1 + \delta_3(\tau))$ with $|\delta_1(\tau)| \ll 1$ and $|\delta_3(\tau)| \ll 1$ perturbations to the steady-state (equilibrium) values $\eta_{1,\infty}$ and $\eta_{3,\infty}$ and linearize them. This produces a set of equations that can be put in the form of the matrix equation $d\delta/d\tau = -\mathbf{M} \cdot \delta$, where $\delta \equiv (\delta_1, \delta_3)^T$ is the perturbation vector and the kinetic matrix $\mathbf{M}$ becomes

$$\mathbf{M} \equiv \begin{bmatrix} \frac{1}{\eta_{s,\infty}} + \frac{1}{q_1\eta_{1,\infty}} & \frac{\eta_{3,\infty}}{\eta_{1,\infty}\eta_{s,\infty}} \\ \frac{\eta_{1,\infty}}{\eta_{3,\infty}\eta_{s,\infty}} & \Gamma \left(\frac{1}{\eta_{s,\infty}} + \frac{1}{q_3\eta_{3,\infty}} \right) \end{bmatrix} \quad (\text{A1})$$

with $\eta_{s,\infty} = 1 - \eta_{1,\infty} - \eta_{3,\infty}$ the equilibrium fraction of protein subunits free in solution. To find the eigenvalues $\lambda_{\pm}$ of the matrix $\mathbf{M}$, we write $\det(\mathbf{M} - \lambda_{\pm}\mathbf{I}) = 0$ with $\mathbf{I}$ the identity matrix, and obtain

$$\lambda_{\pm} = \frac{1}{2}a \pm \frac{1}{2}\sqrt{a^2 - 4b} \quad (\text{A2})$$

with

$$a \equiv \frac{1}{\eta_{s,\infty}} + \frac{1}{q_1\eta_{1,\infty}} + \Gamma \left(\frac{1}{\eta_{s,\infty}} + \frac{1}{q_3\eta_{3,\infty}} \right) \quad (\text{A3})$$

and

$$b \equiv \frac{\Gamma}{\eta_{s,\infty}} \left[\frac{1}{q_1\eta_{1,\infty}} + \frac{1}{q_3\eta_{3,\infty}} + \frac{\eta_{s,\infty}}{q_1\eta_{1,\infty}q_3\eta_{3,\infty}} \right]. \quad (\text{A4})$$

In the limit where $4b/a^2 \ll 1$, the eigenvalues representing the fundamental relaxation rates reduce to

$$\lambda_{+} \sim a \times \left(1 - \frac{b}{a^2} \right) \quad (\text{A5})$$

and

$$\lambda_{-} \sim a \times \frac{b}{a^2} \quad (\text{A6})$$

up to linear order in b/a^2 . Notice that the leading order $\lambda_{-} \sim \lambda_{+}b/a^2$, implying that $\lambda_{-} \ll \lambda_{+}$ and there must be a strong separation of time scales. Under conditions where $q_1\eta_{1,\infty} \gg \eta_{s,\infty}$ and $q_3\eta_{3,\infty} \gg \eta_{s,\infty}$, these expressions simplify to Eqs. (13) and (14). For these conditions to hold, we must have $q_1\eta_{1,\infty} \gg 1$ and $q_3\eta_{3,\infty} \gg 1$: the equilibrium (final) fraction of proteins in both types of capsids cannot be smaller than the reciprocal of their aggregation numbers.

Obviously, for the rates to remain real numbers, we must insist that $4b \leq a^2$. For $\Gamma \rightarrow 0$ and $\Gamma \rightarrow \infty$, this is easily verified, but for arbitrary values, this is not so trivial, noting that $\eta_{1,\infty}$ and $\eta_{3,\infty}$ are not independent [see Eq. (6)]. Figure 4, showing $\lambda_{\pm}$ for $\Gamma = 1$, suggests this is always the case. Finally, if $4b \rightarrow a^2$, the rates approach each other and become equal to

$$\lambda_{\pm} \sim \frac{1}{2}a, \quad (\text{A7})$$

indicating, in this particular case, the presence of a single time scale.

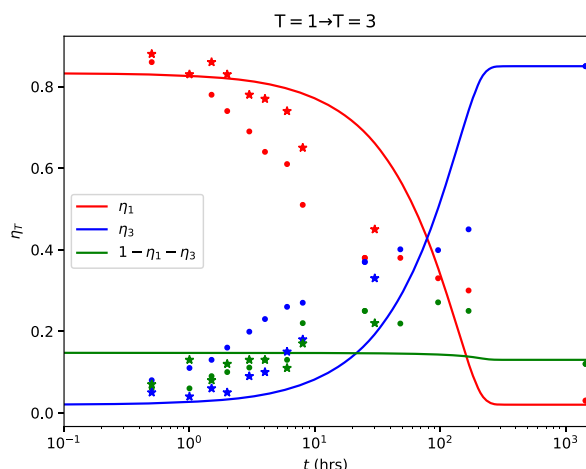


FIG. 11. Fraction of protein in various species η_T as a function of time t in hours. Symbols: results from two datasets where, following a quench, $T = 1$ capsids convert into $T = 3$ capsids. For two months, the measurements of Timmermans *et al.*⁴⁸ are shown. The red, blue, and green lines indicate the fractions of $T = 1$, $T = 3$, and free subunits. See the text in Appendix B.

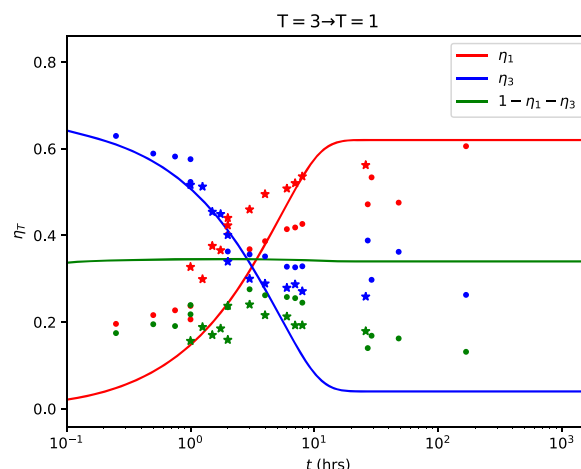


FIG. 12. Fraction of protein in various species η_T as a function of time t in hours. Symbols: results from two datasets where, following a quench, $T = 3$ capsids convert into $T = 1$ capsids. For two months, the measurements of Timmermans *et al.*⁴⁸ are shown. The red, blue, and green lines indicate the fractions of $T = 1$, $T = 3$, and free subunits. See also the text in Appendix B.

The solution for $\delta(\tau)$, can be written in terms of a linear combination of the (unnormalized) eigenvectors $\mathbf{v}_{\pm}$ associated with the eigenvalues $\lambda_{\pm}$,

$$\delta(\tau) = c_+ \mathbf{v}_+ \exp(-\lambda_+ \tau) + c_- \mathbf{v}_- \exp(-\lambda_- \tau), \quad (\text{A8})$$

where

$$\mathbf{v}_{\pm} = \begin{pmatrix} \frac{\eta_{3,\infty} \eta_{1,\infty}^{-1} \eta_{s,\infty}^{-1}}{\eta_{s,\infty}^{-1} + q_1^{-1} \eta_{1,\infty}^{-1} - \lambda_{\pm}} \\ -1 \end{pmatrix} \quad (\text{A9})$$

and c_+ and c_- are constants that are fixed by the initial conditions $\delta(0) = c_+ \mathbf{v}_+ + c_- \mathbf{v}_-$, noting that $\delta_T(0) = (\eta_T(0)/\eta_{T,\infty}) - 1$. Notice also that $\mathbf{v}_{\pm}$ is not orthogonal unless $\Gamma = 1$ and $\eta_{1,\infty} = \eta_{3,\infty}$, in which case the matrix $\mathbf{M}$ is symmetric. If $4b = a^2$, there is only one eigenvector.

APPENDIX B: COMPLETE CAPSID CONVERSIONS

We made an attempt to fit the experimental data precisely according to our theory, focusing on achieving an exact match between the concentrations of $T = 1$ and $T = 3$ after two months. To account for the experimental conditions where only one species is expected to survive at the end, we assigned greater significance to the end points during the data fitting process. Figures 11 and 12 clearly indicate that obtaining a satisfactory fit for the experimental data is not feasible under these conditions.

Figure 11 shows the experimental data illustrating the conversion of a solution primarily composed of $T = 1$ particles to a solution where the dominant species are the $T = 3$ particles after $t = 1440$ h. In line with the main text, the symbols represent two sets of experimental data, while the solid lines represent the fitted curves. Our objective is to determine the best-fitting line that passes through the

TABLE II. Numerical values of the model parameters obtained from fitting the relaxational model to the experimental data for the complete two month conversions of $T = 1 \rightarrow T = 3$ and $T = 3 \rightarrow T = 1$ capsids with $d\tau = 0.01$. Γ_1 and Γ_3 are the relaxation rates associated with the assembly and disassembly of the $T = 1$ and $T = 3$ capsids, and $\Gamma = \Gamma_3/\Gamma_1$ is their ratio. The dimensionless free energies of binding of coat protein subunits are g_1 and g_3 , and $\Delta g = g_3 - g_1$ is their difference. See also the main text.

Conversion	Γ_1 (h^{-1})	Γ_3 (h^{-1})	Γ	g_1	g_3	Δg
$T = 1 \rightarrow T = 3$	39.0	0.04	0.001	-11.4	-11.8	-0.4
$T = 3 \rightarrow T = 1$	1.8	7.8	4.4	-10.5	-10.8	-0.3

last final experimental data point. The initial conditions are set to $\eta_1(0) = 0.86$ and $\eta_3(0) = 0.02$, and the equilibrium fractions are set at $\eta_{1,\infty} = 0.02$ and $\eta_{3,\infty} = 0.85$. Using these equilibrium fractions, we derive binding energies of $g_1 = -11.4$ and $g_3 = -11.8$, indicating a difference of $\Delta g = -0.4$. The corresponding rate constants are $\Gamma_1 = 39.0$ and $\Gamma_3 = 0.04 \text{ h}^{-1}$, resulting in $\Gamma = 0.001$. The agreement between theory and experiment is clearly not ideal, with an r^2 value of 0.61. Table II presents the rate constants and binding free energies obtained from our data fitting analysis of the interconversion between $T = 1$ and $T = 3$.

Figure 12 shows the experimental data illustrating the conversion of a solution primarily composed of $T = 3$ particles to a solution where the dominant species are the $T = 1$ particles after $t = 1608$ h. The initial conditions are set to $\eta_1(0) = 0.01$ and $\eta_3(0) = 0.72$, and the equilibrium fractions are set at $\eta_{1,\infty} = 0.62$ and $\eta_{3,\infty} = 0.04$. Using these equilibrium fractions, we derive binding energies of $g_1 = -10.5$ and $g_3 = -10.8$, indicating a difference of $\Delta g = -0.3$. The corresponding rate constants are $\Gamma_1 = 1.8$ and $\Gamma_3 = 7.8 \text{ h}^{-1}$, resulting in $\Gamma = 4.4$. The agreement between theory and experiment is again not ideal, with an r^2 value of 0.425. Table II presents the rate

constants and binding free energies obtained from our data fitting analysis of the interconversion between $T = 3$ and $T = 1$.

REFERENCES

- ¹R. F. Bruinsma, G. J. L. Wuite, and W. H. Roos, "Physics of viral dynamics," *Nat. Rev. Phys.* **3**, 76–91 (2021).
- ²M. F. Hagan, "Modeling viral capsid assembly," *Adv. Chem. Phys.* **155**, 1–68 (2014).
- ³E. C. Dykeman, P. G. Stockley, and R. Twarock, "Solving a Levinthal's paradox for virus assembly identifies a unique antiviral strategy," *Proc. Natl. Acad. Sci. U. S. A.* **111**, 5361–5366 (2014).
- ⁴B. Rabe, M. Delaleau, A. Bischof, M. Foss, I. Sominskaya, P. Pumpens, C. Cazenave, M. Castroviejo, and M. Kann, "Nuclear entry of hepatitis B virus capsids involves disintegration to protein dimers followed by nuclear reassociation to capsids," *PLoS Pathog.* **5**, e1000563 (2009).
- ⁵S. Li and R. Zandi, "Biophysical modeling of SARS-CoV-2 assembly: Genome condensation and budding," *Viruses* **14**, 2089 (2022).
- ⁶J. Bancroft and E. Hiebert, "Formation of an infectious nucleoprotein from protein and nucleic acid isolated from a small spherical virus," *Virology* **32**, 354–356 (1967).
- ⁷J. B. Bancroft, "The self-assembly of spherical plant viruses," *Adv. Virus Res.* **16**, 99–134 (1970).
- ⁸K. W. Adolph and P. J. G. Butler, "Studies on the assembly of a spherical plant virus. I. States of aggregation of the isolated protein," *J. Mol. Biol.* **88**, 327–341 (1974).
- ⁹N. L. Goicochea, M. De, V. M. Rotello, S. Mukhopadhyay, and B. Dragnea, "Core-like particles of an enveloped animal virus can self-assemble efficiently on artificial templates," *Nano Lett.* **7**, 2281–2290 (2007).
- ¹⁰Q. Liu, A. Shaikat, D. Kyllönen, and M. A. Kostiaainen, "Polyelectrolyte encapsulation and confinement within protein cage-inspired nanocompartments," *Pharmaceutics* **13**, 1551 (2021).
- ¹¹R. F. Garmann, M. Comas-Garcia, C. M. Knobler, and W. M. Gelbart, "Physical principles in the self-assembly of a simple spherical virus," *Acc. Chem. Res.* **49**, 48–55 (2016).
- ¹²L. He, Z. Porterfield, P. van der Schoot, A. Zlotnick, and B. Dragnea, "Hepatitis virus capsid polymorph stability depends on encapsulated cargo size," *ACS Nano* **7**, 8447–8454 (2013).
- ¹³H. S. Savithri and J. W. Erickson, "The self-assembly of the cowpea strain of southern bean mosaic virus: Formation of $T = 1$ and $T = 3$ nucleoprotein particles," *Virology* **126**, 328–335 (1983).
- ¹⁴K. Bond, I. B. Tsvetkova, J. C.-Y. Wang, M. F. Jarrold, and B. Dragnea, "Virus assembly pathways: Straying away but not too far," *Small* **16**, 2004475 (2020).
- ¹⁵S. Panahandeh, S. Li, B. Dragnea, and R. Zandi, "Virus assembly pathways inside a host cell," *ACS Nano* **16**, 317–327 (2022).
- ¹⁶S. Paquay, H. Kusumaatmaja, D. J. Wales, R. Zandi, and P. van der Schoot, "Energetically favoured defects in dense packings of particles on spherical surfaces," *Soft Matter* **12**, 5708–5717 (2016).
- ¹⁷S. Panahandeh, S. Li, and R. Zandi, "The equilibrium structure of self-assembled protein nano-cages," *Nanoscale* **10**, 22802–22809 (2018).
- ¹⁸A. Zlotnick, N. Cheng, J. F. Conway, F. P. Booy, A. C. Steven, S. J. Stahl, and P. T. Wingfield, "Dimorphism of hepatitis B virus capsids is strongly influenced by the C-terminus of the capsid protein," *Biochemistry* **35**, 7412–7421 (1996).
- ¹⁹X. Sun, D. Li, Z. Wang, Q. Liu, Y. Wei, and T. Liu, "A dimorphism shift of hepatitis B virus capsids in response to ionic conditions," *Nanoscale* **10**, 16984–16989 (2018).
- ²⁰R. Zandi and P. van der Schoot, "Size regulation of ss-RNA viruses," *Biophys. J.* **96**, 9–20 (2009).
- ²¹P. van der Schoot and R. Zandi, "Impact of the topology of viral RNAs on their encapsulation by virus coat proteins," *J. Biol. Phys.* **39**, 289–299 (2013).
- ²²P. Moerman, P. van der Schoot, and W. Kegel, "Kinetics vs thermodynamics in virus capsid polymorphism," *J. Phys. Chem. B* **120**, 6003–6009 (2016).
- ²³K. W. Adolph and P. J. G. Butler, "Assembly of a spherical plant virus," *Philos. Trans. R. Soc. London, Ser. B* **276**, 113–122 (1976).
- ²⁴L. Lavelle, M. Gingery, M. Phillips, W. M. Gelbart, C. M. Knobler, R. D. Cadena-Nava, J. R. Vega-Acosta, L. A. Pinedo-Torres, and J. Ruiz-Garcia, "Phase diagram of self-assembled viral capsid protein polymorphs," *J. Phys. Chem. B* **113**, 3813–3819 (2009).
- ²⁵S. Li, G. Erdemci-Tandogan, J. Wagner, P. van der Schoot, and R. Zandi, "Impact of a nonuniform charge distribution on virus assembly," *Phys. Rev. E* **96**, 022401 (2017).
- ²⁶Y. Dong, S. Li, and R. Zandi, "Effect of the charge distribution of virus coat proteins on the length of packaged RNAs," *Phys. Rev. E* **102**, 062423 (2020).
- ²⁷P. Ni, Z. Wang, X. Ma, N. C. Das, P. Sokol, W. Chiu, B. Dragnea, M. Hagan, and C. C. Kao, "An examination of the electrostatic interactions between the N-terminal tail of the Brome Mosaic Virus coat protein and encapsidated RNAs," *J. Mol. Biol.* **419**, 284–300 (2012).
- ²⁸P. Ceres and A. Zlotnick, "Weak protein-protein interactions are sufficient to drive assembly of hepatitis B virus capsids," *Biochemistry* **41**, 11525–11531 (2002).
- ²⁹R. F. Bruinsma, W. M. Gelbart, D. Reguera, J. Rudnick, and R. Zandi, "Viral self-assembly as a thermodynamic process," *Phys. Rev. Lett.* **90**, 248101–248105 (2003).
- ³⁰D. Caspar, "Assembly and stability of the tobacco mosaic virus particle," *Adv. Protein Chem.* **18**, 37–121 (1964).
- ³¹W. Chiu, R. Burnett, and R. Garcea, *Structural Biology of Viruses* (Oxford University Press, 1997).
- ³²A. McPherson, "Micelle formation and crystallization as paradigms for virus assembly," *BioEssays* **27**, 447–458 (2005).
- ³³A. Zlotnick, "To build a virus capsid. An equilibrium model of the self assembly of polyhedral protein complexes," *J. Mol. Biol.* **241**, 59–67 (1994).
- ³⁴T. Keef, C. Micheletti, and R. Twarock, "Master equation approach to the assembly of viral capsids," *J. Theor. Biol.* **242**, 713–721 (2006).
- ³⁵R. Zandi, P. van der Schoot, D. Reguera, W. Kegel, and H. Reiss, "Classical nucleation theory of virus capsids," *Biophys. J.* **90**, 1939–1948 (2006).
- ³⁶C. Berthet-Colominas, M. Cuillel, M. H. J. Koch, P. Vachette, and B. Jacrot, "Kinetic study of the self-assembly of brome mosaic virus capsid," *Eur. Biophys. J.* **15**, 159–168 (1987).
- ³⁷D. C. Rapaport, "Molecular dynamics simulation: A tool for exploration and discovery using simple models," *J. Phys.: Condens. Matter* **26**, 503104 (2014).
- ³⁸R. F. Garmann, A. M. Goldfain, and V. N. Manoharan, "Measurements of the self-assembly kinetics of individual viral capsids around their RNA genome," *Proc. Natl. Acad. Sci. U. S. A.* **116**, 22485–22490 (2019).
- ³⁹R. Zandi, B. Dragnea, A. Travesset, and R. Podgornik, "On virus growth and form," *Phys. Rep.* **847**, 1–102 (2020).
- ⁴⁰D. Law-Hine, A. K. Sahoo, V. Baillieux, M. Zeghal, S. Prevost, P. K. Maiti, S. Bressanelli, D. Constantin, and G. Tresset, "Reconstruction of the disassembly pathway of an icosahedral viral capsid and shape determination of two successive intermediates," *J. Phys. Chem. Lett.* **6**, 3471–3476 (2015).
- ⁴¹A. Šiber and R. Podgornik, "Nonspecific interactions in spontaneous assembly of empty vs functional single-stranded RNA viruses," *Phys. Rev. E* **78**, 051915 (2008).
- ⁴²M. Chevreuil, D. Law-Hine, J. Chen, S. Bressanelli, S. Combet, D. Constantin, J. Degrouard, J. Möller, M. Zeghal, and G. Tresset, "Nonequilibrium self-assembly dynamics of icosahedral viral capsids packaging genome or polyelectrolyte," *Nat. Commun.* **9**, 3071 (2018).
- ⁴³O. M. Elrad and M. F. Hagan, "Encapsulation of a polymer by an icosahedral virus," *Phys. Biol.* **7**, 045003 (2010).
- ⁴⁴S. Panahandeh, S. Li, L. Marichal, R. Leite Rubim, G. Tresset, and R. Zandi, "How a virus circumvents energy barriers to form symmetric shells," *ACS Nano* **14**, 3170–3180 (2020).
- ⁴⁵M. Chevreuil, L. Lecoq, S. Wang, L. Gargowitsch, N. Nhiri, E. Jacquet, T. Zinn, S. Fieulaine, S. Bressanelli, and G. Tresset, "Nonsymmetrical dynamics of the HBV capsid assembly and disassembly evidenced by their transient species," *J. Phys. Chem. B* **124**, 9987–9995 (2020).
- ⁴⁶S. Singh and A. Zlotnick, "Observed hysteresis of virus capsid disassembly is implicit in kinetic models of assembly," *J. Biol. Chem.* **278**, 18249–18255 (2003).
- ⁴⁷Z. Ghaemi, M. Gruebele, and E. Tajkhorshid, "Molecular mechanism of capsid disassembly in hepatitis B virus," *Proc. Natl. Acad. Sci. U. S. A.* **118**, e2102530118 (2021).

- ⁴⁸S. B. P. E. Timmermans, A. Ramezani, T. Montalvo, M. Nguyen, P. van der Schoot, J. C. M. van Hest, and R. Zandi, "The dynamics of viruslike capsid assembly and disassembly," *J. Am. Chem. Soc.* **144**, 12608–12612 (2022).
- ⁴⁹J. R. Vega-Acosta, R. D. Cadena-Nava, W. M. Gelbart, C. M. Knobler, and J. Ruiz-García, "Electrophoretic mobilities of a viral capsid, its capsid protein, and their relation to viral assembly," *J. Phys. Chem. B* **118**, 1984–1989 (2014).
- ⁵⁰P. C. Hohenberg and B. I. Halperin, "Theory of dynamic critical phenomena," *Rev. Mod. Phys.* **49**, 435–479 (1977).
- ⁵¹P. van der Schoot and R. Zandi, "Kinetic theory of virus capsid assembly," *Phys. Biol.* **4**, 296–304 (2007).
- ⁵²D. C. Rappaport, "Role of reversibility in viral capsid growth: A paradigm for self-assembly," *Phys. Rev. Lett.* **101**, 186101 (2008).
- ⁵³Y. Hu, R. Zandi, A. Anavitarte, C. M. Knobler, and W. M. Gelbart, "Packaging of a polymer by a viral capsid: The interplay between polymer length and capsid size," *Biophys. J.* **94**, 1428–1436 (2008).
- ⁵⁴W. K. Kegel and P. van der Schoot, "Competing hydrophobic and screened-coulomb interactions in hepatitis B virus capsid assembly," *Biophys. J.* **86**, 3905–3913 (2004).
- ⁵⁵P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, Í. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, "SciPy 1.0: Fundamental algorithms for scientific computing in Python," *Nat. Methods* **17**, 261–272 (2020).
- ⁵⁶P. Lerman, "Fitting segmented regression models by grid search," *J. R. Stat. Soc. Ser. C: Appl. Stat.* **29**, 77–84 (1980).
- ⁵⁷D. C. Montgomery, *Design and Analysis of Experiments*, 8th ed. (John Wiley & Sons, Inc., Hoboken, NJ, 2013).
- ⁵⁸P. Rein ten Wolde and D. Frenkel, "Homogeneous nucleation and the Ostwald step rule," *Phys. Chem. Chem. Phys.* **1**, 2191–2196 (1999).
- ⁵⁹D. Kashchiev, "Classical nucleation theory approach to two-step nucleation of crystals," *J. Cryst. Growth* **530**, 125300 (2020).
- ⁶⁰A. Y. Morozov, R. F. Bruinsma, and J. Rudnick, "Assembly of viruses and the pseudo-law of mass action," *J. Chem. Phys.* **131**, 155101 (2009).
- ⁶¹L. Schoonen, R. J. M. Maas, R. J. M. Nolte, and J. C. M. van Hest, "Expansion of the assembly of cowpea chlorotic mottle virus towards non-native and physiological conditions," *Tetrahedron* **73**, 4968–4971 (2017).
- ⁶²E. C. Dykeman, P. G. Stockley, and R. Twarock, "Dynamic allostery controls coat protein conformer switching during MS2 phage assembly," *J. Mol. Biol.* **395**, 916–923 (2010).