PNAS RESEARCH ARTICLE | BIOPHYSICS AND COMPUTATIONAL BIOLOGY OPEN ACCESS

Single-particle studies of the effects of RNA–protein interactions on the self-assembly of RNA virus particles

Rees F. Garmann[a,b,c,1,2], Aaron M. Goldfain[a,d,1], Cheylene R. Tanimoto[e,1], Christian E. Beren[e,f], Fernando F. Vasquez[b], Daniel A. Villarreal[b], Charles M. Knobler[e], William M. Gelbart[e,g,h], and Vinothan N. Manoharan[a,i]

Edited by Owen Pornillos, University of Virginia School of Medicine; received April 24, 2022; accepted August 23, 2022 by Editorial Board Member Michael F. Summers

Understanding the pathways by which simple RNA viruses self-assemble from their coat proteins and RNA is of practical and fundamental interest. Although RNA–protein interactions are thought to play a critical role in the assembly, our understanding of their effects is limited because the assembly process is difficult to observe directly. We address this problem by using interferometric scattering microscopy, a sensitive optical technique with high dynamic range, to follow the in vitro assembly kinetics of more than 500 individual particles of brome mosaic virus (BMV)—for which RNA–protein interactions can be controlled by varying the ionic strength of the buffer. We find that when RNA–protein interactions are weak, BMV assembles by a nucleation-and-growth pathway in which a small cluster of RNA-bound proteins must exceed a critical size before additional proteins can bind. As the strength of RNA–protein interactions increases, the nucleation time becomes shorter and more narrowly distributed, but the time to grow a capsid after nucleation is largely unaffected. These results suggest that the nucleation rate is controlled by RNA–protein interactions, while the growth process is driven less by RNA–protein interactions and more by protein–protein interactions and intraprotein forces. The nucleated pathway observed with the plant virus BMV is strikingly similar to that previously observed with bacteriophage MS2, a phylogenetically distinct virus with a different host kingdom. These results raise the possibility that nucleated assembly pathways might be common to other RNA viruses.

RNA virus | viral self-assembly | kinetics | iSCAT | single-particle analyses

Significance

RNA viruses first inspired the term "self-assembly." Yet much is still not understood about how even the simplest such viruses assemble or if different viruses assemble in similar ways. Theoretical models suggest many possible assembly pathways, with many different roles for RNA, but until recently measuring these pathways has not been possible. We use a sensitive microscopy technique to follow the assembly of individual particles of brome mosaic virus (BMV), a plant virus. We find evidence of an RNA-mediated nucleation-and-growth pathway that is strikingly similar to that of MS2, a bacterial virus. The last common ancestor of BMV and MS2 existed only in ancient times, suggesting that their assembly pathway might be evolutionarily conserved and that other viruses might follow a similar pathway.

Since the 1950s, the question of how RNA viruses self-assemble has inspired theoretical and experimental work in many fields of basic and applied science (1–5). Simple RNA viruses, which consist of a single-stranded RNA genome inside an ordered capsid made up of multiple copies of a single protein (Fig. 1*A*), have served as model systems for studying the physical principles of structural virology involving virus particles of all shapes and sizes (1, 2, 6, 7). However, the mechanisms and pathways by which these viruses assemble into the correct structure, while avoiding the many possible malformed structures, are not yet understood.

Although many different RNA viruses self-assemble (8–10), our interest is in comparing the assembly of virus-like particles from two well-studied virus families: *Bromoviridae*, a family of plant-infecting viruses that includes brome mosaic virus (BMV) and cowpea chlorotic mottle virus (CCMV), and *Fiersviridae* (previously *Leviviridae*), a family of bacteria-infecting viruses that includes MS2 and Qβ. These families are as distinct phylogenetically as any two RNA virus families can be, having a last common ancestor that is thought to predate the emergence of eukaryotic cells (11). Accordingly, there are many well-established physical and biological differences among viruses in these families and virus-like particles derived from them. Yet the four most studied members—BMV, CCMV, MS2, and Qβ—do have some structural commonalities: They have icosahedral capsids with a triangulation number (T) of 3 (2), they have no lipid envelope, and each capsid surrounds approximately 3,000 to 4,000 nucleotides of single-stranded RNA.

The assembly of such structures is a nontrivial process. Identical coat proteins must adopt nonequivalent positions to make a $T = 3$ capsid, with some arranging in pentagonal configurations and others in hexagonal configurations (2, 7, 12). Furthermore, these configurations must form in the correct proportions and positions for the capsid to close. Despite these challenges, assembly of virus-like particles of CCMV (13–15), BMV (14, 15), and MS2 (16) occurs in high yield even in vitro and in the absence of host-cell factors. The ability of viruses to avoid the many possible metastable states en route to complete assembly has been likened to the Levinthal paradox of protein folding (17, 18).

But unlike proteins, RNA viruses have a template for assembly: their own RNA. Current theoretical models of RNA virus self-assembly posit markedly different roles for the RNA, depending on the relative strengths of RNA–protein and protein–protein

Author contributions: R.F.G., A.M.G., C.R.T., C.E.B., F.F.V., D.A.V., C.M.K., W.M.G., and V.N.M. designed research; R.F.G., A.M.G., C.R.T., C.E.B., F.F.V., and D.A.V. performed experiments; R.F.G., A.M.G., C.R.T., C.E.B., F.F.V., D.A.V., C.M.K., W.M.G., and V.N.M. analyzed data; R.F.G., C.M.K., W.M.G., and V.N.M. supervised the project; and R.F.G., A.M.G., C.R.T., C.E.B., F.F.V., D.A.V., C.M.K., W.M.G., and V.N.M. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission. O.P. is a guest editor invited by the Editorial Board.

Copyright © 2022 the Author(s). Published by PNAS. This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

[1]R.F.G., A.M.G., and C.R.T. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: rgarmann@sdsu.edu.

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2206292119/-/DCSupplemental.

Published September 19, 2022.

Downloaded from https://www.pnas.org by NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (NIST) on September 22, 2022 from IP address 129.6.219.63.

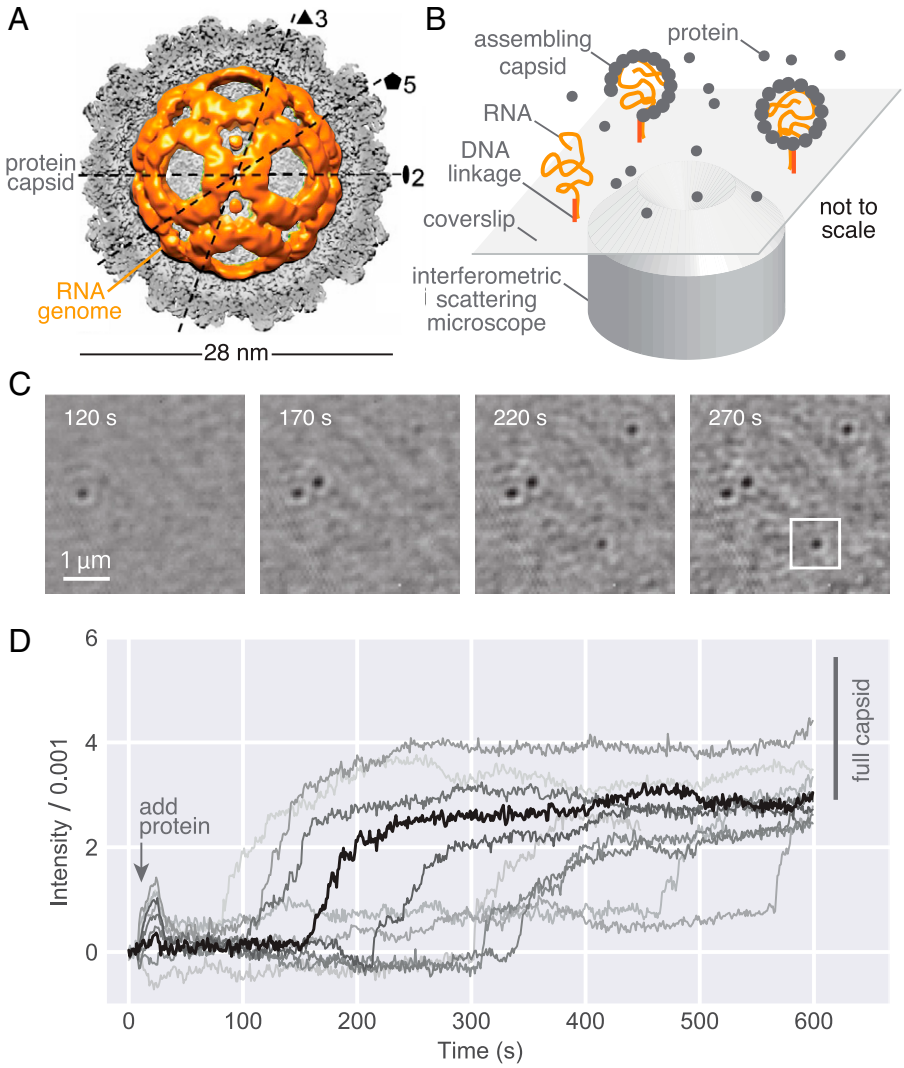


Fig. 1. Overview of the system and the measurement. (*A*) A 3-dimensional model of BMV reconstructed from cryoelectron microscopy data (51) shows the protein capsid (gray) surrounding the RNA (gold). The model reveals most of the icosahedral capsid but only a small portion of the RNA, the rest of which adopts a disordered arrangement within the capsid. (*B*) A cartoon of the experiment shows viral coat proteins assembling around RNA strands that are tethered by DNA linkages to the surface of a functionalized glass coverslip. (*C*) The assembling proteins are imaged at 1,000 Hz for 600 s using iSCAT microscopy. Each dark spot that appears in the images corresponds to proteins bound to an individual RNA strand. The darkness, or intensity, of each spot is proportional to the number of proteins bound to that RNA. The displayed images are the average of 1,000 consecutive frames. (*D*) Traces of the intensity as a function of time (1,000-frame moving average) reveal the assembly kinetics for each particle. Experimental conditions are 0.135 μmol/L protein and 250 mmol/L NaCl. The initial spike in intensity present in many of the traces is associated with vibrations introduced into the system as the coat protein is injected. The thick, black trace corresponds to the boxed particle in (*C*). We compare the final intensities of the traces to the estimated intensity range of full capsids, which is shown as a vertical bar to the right of the traces.

interactions, sequence-dependent RNA–protein interactions, RNA-mediated protein–protein interactions, and several other factors (19). Although specific interactions between RNA substructures and coat proteins have been hypothesized to help the virus avoid malformed configurations (18), viruses from different families differ greatly in their RNA structures and RNA–protein interactions. It is therefore unclear whether there are common features of the assembly process for different $T = 3$ viruses or if there are distinct assembly pathways that depend on RNA–protein interactions.

Recent measurements of assembly kinetics suggest the latter: that the assembly of viruses from different families follows different pathways. Fluorescence correlation spectroscopy experiments (20, 21) of the kinetics of binding of MS2 coat protein and RNA indicate that assembly starts with a small cluster of RNA-bound proteins that trigger a change in the hydrodynamic radius of the RNA. In contrast, cryoelectron microscopy (22) and small-angle X-ray scattering (23) experiments of the assembly of the CCMV coat protein and RNA show that disordered RNA–protein complexes formed at neutral pH anneal over several thousand seconds into well-formed capsids when the pH drops below 6.

But because these experiments involve different assembly conditions and different measurement techniques, their outcomes might not reflect fundamental differences in the assembly pathways of these viruses but rather technical differences in the methods and protocols used to study them. Furthermore, most of the techniques that have been used do not measure the assembly process directly at the scale of individual particles because—one way or the other—they involve averaging over many particles. Such averaging can obscure the mechanisms and pathways that underpin stochastic assembly processes like viral assembly, in which each individual particle can follow its own unique sequence of intermediate states. Thus, it remains an open question whether a common assembly pathway might exist between these viruses.

Downloaded from https://www.pnas.org by NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (NIST) on September 22, 2022 from IP address 129.6.219.63.

We recently demonstrated that interferometric scattering (iSCAT) microscopy (24) can resolve the assembly kinetics of individual virus-like particles (25), providing a method to directly measure and compare the assembly pathways of different viruses. To perform the iSCAT experiment, we first tether viral RNA molecules to the surface of a functionalized glass coverslip under the desired buffer conditions (26) (Fig. 1*B*). Next, we begin collecting iSCAT images of the RNA-decorated coverslip as we inject viral coat proteins at the desired concentration and in the appropriate buffer. As the proteins bind to the surface-tethered RNA, dark spots appear in the iSCAT images (Fig. 1*C*). Subtracting the intensity associated with the RNA then yields images in which the intensity of each dark spot is proportional to the number of proteins that have accrued onto each individual RNA. Previous measurements by Young and coworkers (27) show that the iSCAT intensities of protein assemblies are, in general, linearly proportional to the total mass of the assemblies. Accordingly, in our experiments, plotting the trace of the intensity of a spot as a function of time reveals the assembly kinetics for that particle, and plotting the collection of traces reveals the assembly kinetics for the ensemble of particles (Fig. 1*D*).

In our previous work (25) we examined the assembly of bacteriophage MS2. We found that well-formed capsids could assemble around surface-tethered RNA strands and that the assembly kinetics were consistent with a nucleation-and-growth pathway in which a small cluster of RNA-bound proteins must exceed a critical size before the binding of additional proteins becomes favorable. Despite an apparently small critical nucleus size of only a few coat–protein dimers, we found that MS2 capsids grow monotonically to full or nearly full size with high yield.

Although this previous study highlighted the importance of the RNA in the assembly process, the strong and specific RNA–protein interactions in MS2 (28–30), which are thought to occur at a dozen or so positions on the RNA molecule (31, 32), make it difficult to systematically address the central question of how the RNA affects the pathway. By contrast, the RNA in BMV interacts with the coat proteins primarily through nonspecific electrostatic interactions (33), with the possible exception of a single, specific RNA–protein interaction occurring at the 3′-end of the RNA (34). As a result, the strength of RNA–protein interactions in BMV can be largely tuned by changing the ionic strength of the buffer solution (22, 35, 36). BMV therefore offers not only an interesting comparison to MS2—it is phylogenetically distinct but structurally similar—but also the means to understand the role of RNA–protein interactions.

In the current study, we infer the assembly pathways of BMV from iSCAT measurements under different RNA–protein interaction strengths, allowing us to critically assess of competing models of the assembly process. We follow the assembly trajectories of more than 500 individual virus particles under different assembly conditions, and we correlate the results with the absence and presence of ordered capsids as detected with negative-stain transmission electron microscopy (TEM). We find that BMV can assemble by a nucleation-and-growth process that is qualitatively similar to that of MS2. We show that the strength of RNA–protein interactions strongly affects the nucleation time but only weakly affects the growth time, suggesting that RNA plays a central role in nucleating the viral capsid but a relatively minor role in its growth kinetics. We discuss these observations in the context of recent models and hypotheses of RNA virus self-assembly.

Results

RNA–protein interactions can be controlled by varying ionic strength. We first test whether ionic strength affects the clustering of BMV coat proteins in the absence of RNA. We equilibrate the 20 kDa coat proteins in assembly buffers with different ionic strengths—corresponding to NaCl concentrations of 84, 167, and 250 mmol/L—and measure their mass distributions using iSCAT mass photometry (27). The mass distributions reveal a large peak at 40 kDa, indicating that the coat proteins primarily form dimers (CP_2), with a slight shoulder extending to higher masses, indicating a much smaller population of larger clusters or impurities (Fig. 2*A*). These results are consistent with previous analytical ultracentrifugation measurements (37). Notably, we find that the mass distributions are unchanged by ionic strength (Fig. 2*A*), suggesting that varying NaCl over the range of concentrations tested does not cause coat proteins to change their clustering. Indeed, recent microscale thermophoresis experiments (38) showed that protein–protein interactions between coat proteins in solution are only weakly affected by ionic strength.

To test wheher ionic strength also has a weak effect on the interactions between coat proteins in capsids, we perform differential scanning fluorimetry (DSF) on wild-type BMV particles (39). DSF measures the thermal stability of the virus particles (40), which is related to the strength of lateral interactions between proteins in the capsids. When we weaken protein–protein interactions by raising the pH from 6 to 7 (41), we observe a sharp drop in the thermal stability (Fig. 2*B*), but when we vary ionic strength by varying the concentration of NaCl, we observe no such change (Fig. 2*B*). These results suggest that the interactions between assembled coat proteins, like those between free CP_2, are largely unaffected by ionic strength, at least for the NaCl concentrations tested.

By contrast, we find that the strength of RNA–protein interactions varies markedly with ionic strength. By measuring the yield of RNA–protein binding between 20-nucleotide-long poly-U RNA and BMV CP_2 using a nitrocellulose binding assay (42), we find that increasing the concentration of NaCl decreases the amount of RNA–protein binding (Fig. 2*C*). This result is consistent with binding being driven by electrostatic interactions between the negatively charged phosphate backbone of the RNA and the positively charged N terminus of the coat protein (33). Furthermore, it suggests that we can use ionic strength to specifically modify RNA–protein interactions, largely independently of protein–protein interactions.

Next, we use TEM to test for capsid formation as we vary ionic strength. In assembly reactions involving BMV CP_2 and BMV RNA, we find that 84 mmol/L NaCl leads to heterogeneous assembly products, with some well-formed particles and many malformed particles, whereas 167 mmol/L and 250 mmol/L NaCl give rise to more homogeneous spherical particles with roughly the same size and curvature as wild-type BMV (Fig. 2*D*).

Taken together, these measurements demonstrate that we can control the strength of RNA–protein interactions by varying the ionic strength and that changing these interactions changes the assembly products. These results set the stage for iSCAT single-particle measurements that probe how the assembly pathways might change as RNA–protein interactions are varied.

iSCAT measurements show that BMV assembles by a nucleation-and-growth pathway under certain conditions. We used iSCAT to measure the self-assembly kinetics of individual BMV particles as a function of the RNA–protein interaction strength, as controlled by the ionic strength (84 mmol/L, 167 mmol/L, and 250 mmol/L NaCl) and the protein concentration

Downloaded from https://www.pnas.org by NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (NIST) on September 22, 2022 from IP address 129.6.219.63.

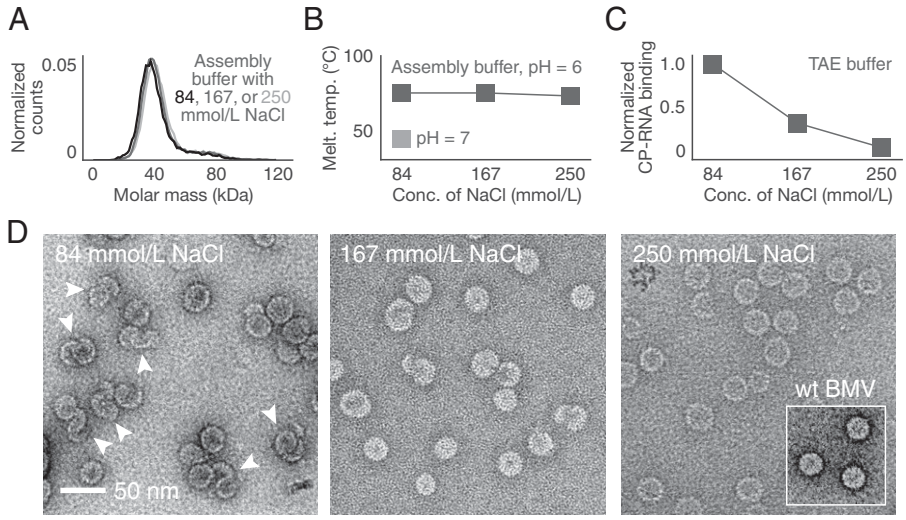


Fig. 2. Ionic strength affects RNA–protein binding and the assembly process but does not affect the clustering of free proteins or the thermal stability of assembled particles. (*A*) iSCAT mass photometry measurements of the mass distribution of BMV coat-protein solutions show 40 kDa dimers as the dominant cluster size for all conditions tested. Measurements are made with 1.7 μg/mL protein in assembly buffer with varying amounts of NaCl. Each of the three overlapping distributions contains at least 15,000 measurements of individual clusters binding to a glass coverslip. The area under each distribution is normalized to a value of 1. (*B*) DSF measurements of the melting temperature of wild-type BMV in assembly buffer at pH 6 with varying NaCl, and at pH 7 with fixed NaCl, show that the capsid stability depends weakly on ionic strength, relative to pH, across the conditions tested. (*C*) Nitrocellulose binding measurements show that the yield of RNA–protein binding depends strongly on ionic strength. Measurements are performed at 4 nmol/L RNA and 0.043 μmol/L CP_2 in TAE buffer with varying amounts of NaCl. The binding yield measurements are normalized with respect to the amount of binding at 84 mmol/L NaCl. (*D*) Uranyl acetate negative-stain TEM images of BMV particles assembled in assembly buffer at varying ionic strengths. We carry out assembly reactions using 7.5 nmol/L BMV RNA1 and 0.86 μmol/L BMV CP_2, and assembly buffers at pH 6 and either 84 mmol/L, 167 mmol/L, or 250 mmol/L NaCl. The 84 mmol/L NaCl sample is notably heterogeneous, presenting many more malformed particles, which are denoted by arrowheads, compared with the higher-NaCl assemblies, which are more monodisperse. Wild-type BMV particles are shown as a reference.

(0.043 μmol/L, 0.135 μmol/L, and 0.427 μmol/L CP_2). In particular, we performed measurements, in duplicate, for the eight combinations of NaCl and protein concentrations indicated in Fig. 3. In total, we measured and analyzed the assembly kinetics of 511 particles, 72 of which are shown in Fig. 3.

We see qualitative differences among these single-particle traces for different RNA–protein interactions, most notably in the time at which each trace begins to increase rapidly (the "start time"). Consider the middle column in Fig. 3, which corresponds to 0.135 μmol/L CP_2. When RNA–protein interactions are strongest (corresponding to 84 mmol/L NaCl), each trace begins increasing immediately after the protein is introduced. In contrast, for weaker RNA–protein interactions (corresponding to 167 mmol/L NaCl), each trace remains at a low intensity for a variable amount of time before increasing, with some traces having start times greater than 100 s at the lowest protein concentration. When RNA–protein interactions are weaker still (corresponding to 250 mmol/L NaCl), the distribution of start times extends to longer values, with some traces having start times as long as 500 s at the intermediate protein concentration. Remarkably, one of the two experiments for 250 mmol/L NaCl did not result in any traces increasing above their initial value over the 600 s experiment.

These experiments allow us to rule out the diffusion-limited accretion of proteins on the RNA as an assembly pathway for at least some of these conditions. If assembly were purely diffusion-limited (43), then we would expect the onset of assembly to vary only with the concentration of protein. Instead, our results show that the start times vary with ionic strength at the same protein concentration.

The variation in start times points to a free-energy barrier to protein accretion on the RNA—in other words, a nucleation barrier (25). In classical nucleation theory, the barrier is associated with an initially unstable cluster—here, of proteins—becoming large enough that subsequent proteins bind favorably (44). Because nucleation is a stochastic process, the time required to form a sufficiently large cluster varies from particle to particle (45, 46). These timescales are reflected in our start-time measurements. The observation that the distribution of start times broadens and shifts to larger values as RNA–protein interactions are weakened (but protein–protein interactions are not significantly changed) reveals that nucleation in BMV is a heterogeneous process, driven at least in part by attractive interactions between the RNA and the assembling proteins.

For the strongest RNA–protein interactions, there is no obvious evidence of a nucleation barrier: At 84 mmol/L NaCl, the traces have nearly identical start times, as shown in the top row of Fig. 3. At these conditions, we see that the final intensities of the traces, a measure of the number of proteins attached to the RNA, increase with increasing protein concentration. Moreover, the final intensities tend to decrease with decreasing RNA–protein interaction strength. This result is qualitatively consistent with our TEM measurements, which show heterogeneous, malformed, larger-than-wild-type particles for the strongest RNA–protein interactions (see structures highlighted by arrowheads in Fig. 2*D*), and increasingly homogeneous, spherical, roughly wild-type-size particles as the RNA–protein interaction strength decreases. Understanding why the assembly products vary with RNA–protein interaction strength requires a more quantitative analysis of the assembly kinetics, and in particular how nucleation rates compare to growth rates.

The shapes of the assembly traces reveal differences between nucleation and growth phases. To extract quantitative information from each recorded trace, we analyze the traces as shown in Fig. 4 to determine three kinetic parameters: the final intensity (I_f), the start time (t_s), and the growth time (τ_g). The values of these parameters for experiments with 0.135 μmol/L CP_2 and varying ionic strength are shown as histograms in Fig. 5.

Downloaded from https://www.pnas.org by NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (NIST) on September 22, 2022 from IP address 129.6.219.63.

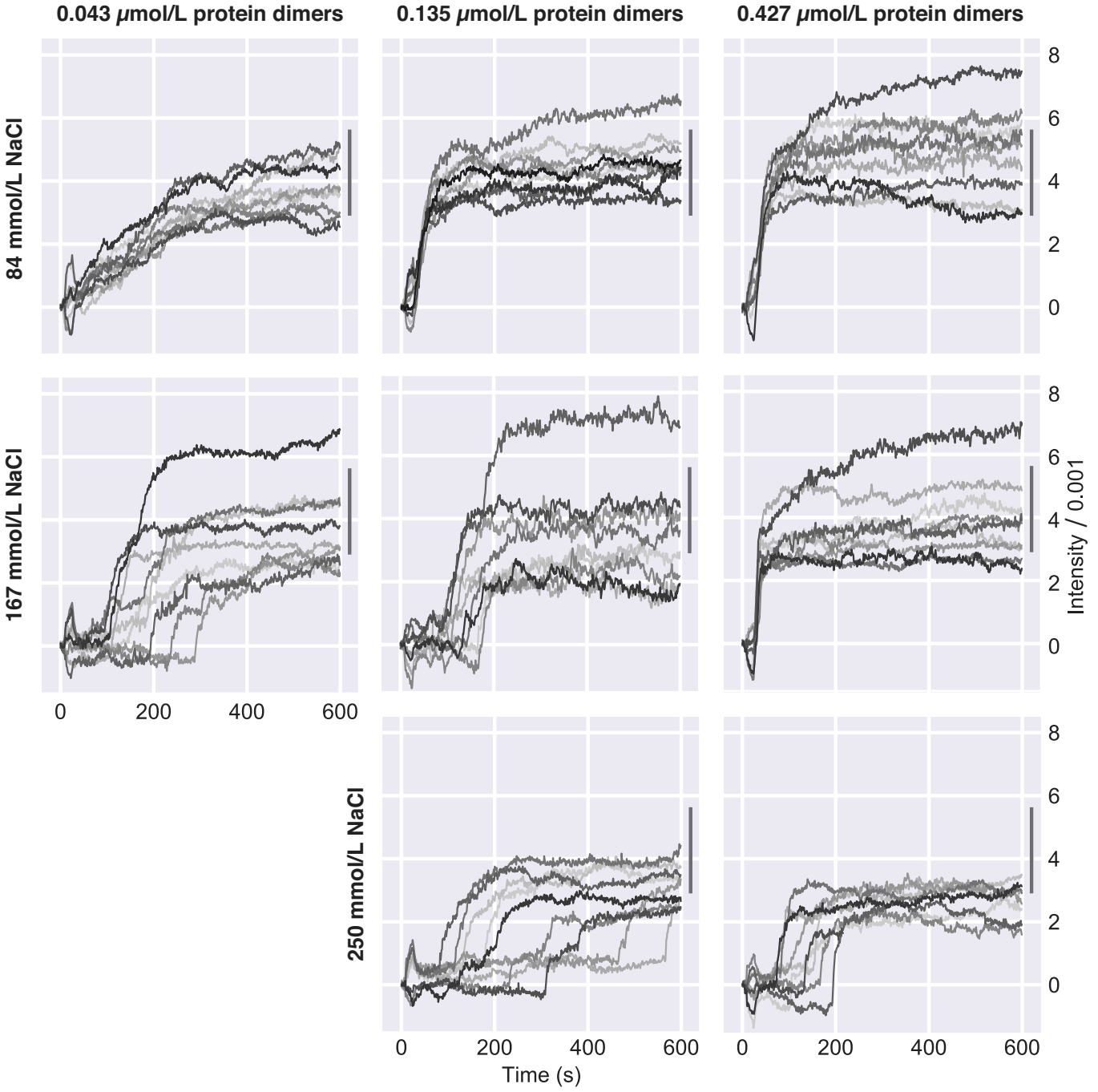


Fig. 3. Single-particle measurements reveal qualitative differences in the assembly kinetics with varying ionic strength and protein concentration. Plots of traces showing intensity as a function of time from iSCAT experiments. Each trace represents the assembly of a single particle consisting of the BMV coat protein and BMV RNA. We show nine randomly selected traces at each condition (the complete set of traces is shown in *SI Appendix*). The bars to the right of the traces show the range of intensities that are consistent with a full capsid.

Consistent with our qualitative analysis of the traces, the start-time histograms show that decreasing the strength of RNA–protein interactions increases the median start times and broadens their spread, as quantified by the median absolute deviation (MAD) (Fig. 5*A*). We find that the median start times increase by more than a factor of 6 and that the MAD of start times increases by more than a factor of 14 when we weaken the RNA–protein interactions by increasing the NaCl concentration from 84 mmol/L to 250 mmol/L (Fig. 5*A*). By contrast, the growth times and final intensities are more narrowly distributed and less affected by changes in RNA–protein interactions, showing an increase by a factor of roughly 1.5 in the median growth times and no increase in the MAD of the growth times (Fig. 5*B*). The median final intensities show a corresponding decrease in particle size by factor of more than 1.5 (Fig. 5*C*).

These results show that RNA–protein interactions primarily affect the nucleation phase of the assembly pathway and only weakly affect the growth phase, if at all. Specifically, stronger RNA–protein interactions reduce the characteristic nucleation time until, at sufficiently large interaction strengths, all particles begin to assemble in apparent synchrony and it is no longer possible to tell if the assembly pathway is nucleated. We discuss this point in more detail in the *Discussion*.

Downloaded from https://www.pnas.org by NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (NIST) on September 22, 2022 from IP address 129.6.219.63.

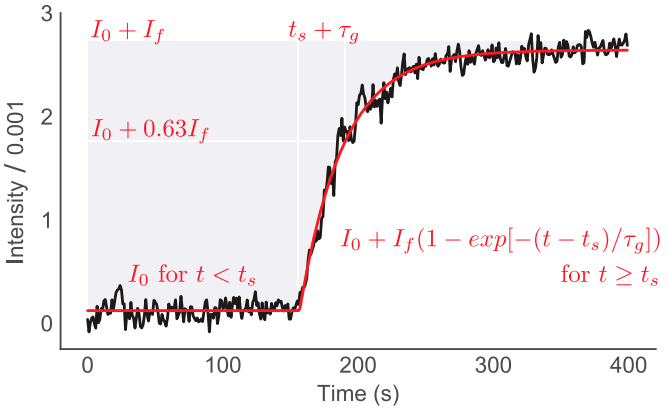


Fig. 4. Extracting quantitative information from the assembly traces: I_f, t_s, and τ_g. An intensity trace of the assembly of a single particle (black curve, corresponding to the boxed particle in Fig. 1*C*) can be fit (red curve) to a piecewise function $I(t) = I_0$ for $t < t_s$; $I(t) = I_0 + I_f\left(1 - \exp[-(t - t_s)/\tau_g]\right)$ for $t \geq t_s$, where I_0 is the initial intensity (which can be offset from zero because of microscope drift), I_f is the final intensity, t_s is the start time (the time at which the trace begins to rise from its initial value), and τ_g is the growth time (the time for the trace to reach (1-1/*e*) = 0.63 of its final value once it has started increasing). The exponential function is not intended to imply a particular assembly mechanism (such as Langmuir adsorption) but rather is chosen because it is a simple function that fits the data well for most traces, allowing us to extract a characteristic growth timescale. In this way, we determine a set of three kinetic parameters, I_f, t_s, and τ_g, for each assembling particle.

To gain further insights into the nucleation kinetics, we examine the effect of protein concentration on the start-time distribution. We plot the MAD of start times, t_{MAD}, for each protein concentration and each RNA–protein interaction strength (corresponding to values of the NaCl concentration) in Fig. 6*A*. In principle, the slope α of the fitted line represents the exponent in a power-law scaling $t_{MAD} \propto c^{-\alpha}$, where c is the protein concentration. We find that for intermediate RNA–protein interaction strength (corresponding to 167 mmol/L NaCl), the measured start times become much more narrowly distributed as the protein concentrations increase, scaling with an exponent $\alpha = 1.5 \pm 0.2$. These results are once again consistent with assembly being a nucleated process—at higher protein concentrations, mass action drives more proteins onto the RNA, favoring larger clusters and shortening the time needed to form a nucleus.

In contrast, for strong RNA–protein interactions (corresponding to 84 mmol/L NaCl), the start times remain narrowly distributed for all protein concentrations tested. Because the experimental uncertainties in our start-time measurements are on the order of seconds, small absolute differences in MAD values are not statistically significant and estimates of the slope α of the fitted line in Fig. 6*A* are less reliable. Thus, the distinction between nucleation and diffusion-limited aggregation becomes blurred. While it is possible that nucleation is occurring on a timescale too fast for us to measure, it is also possible that strengthening RNA–protein interactions qualitatively changes the assembly pathway such that it is no longer nucleated. We discuss these results further in the *Discussion*.

The growth kinetics are much less affected by varying either the protein or NaCl concentration, as shown in Fig. 6*B*. For all but one of the conditions tested, the median growth times differ by a factor of 2.5 at most. Only for 0.043 μmol/L CP_2 and 84 mmol/L NaCl does the median growth time extend to 100 s. And while the median growth times tend to decrease with increasing protein, the amount by which they decrease appears constant for all RNA–protein interaction strengths, as seen by the slopes of the best-fit lines in Fig. 6*B* (slopes are 0.7 ± 0.3 at 84 mmol/L NaCl and 0.5 ± 0.1 at 167 mmol/L NaCl). Thus, while protein concentration seems to have a large effect on the start times, it has a much weaker effect on the growth times. Likewise, while RNA–protein interaction strength has a large effect on the start times, it does not significantly affect the growth times.

These results may explain why larger particles are favored at higher protein concentrations for all RNA–protein interactions tested and at stronger RNA–protein interactions for all protein concentrations tested. These trends can be seen in the heat map shown in Fig. 6*C*. In both cases—increasing protein concentration and increasing RNA–protein interaction strength—the nucleation times decrease relative to the growth times. Multiple nucleation events can therefore happen on the same RNA before any given nucleus has time to grow into a full capsid. Consequently, malformed structures consisting of multiple partially assembled capsids can occur. Indeed, the electron microscope images in Fig. 2*D* show aggregates consisting of multiple capsid-like fragments that have the same curvature as the wild-type

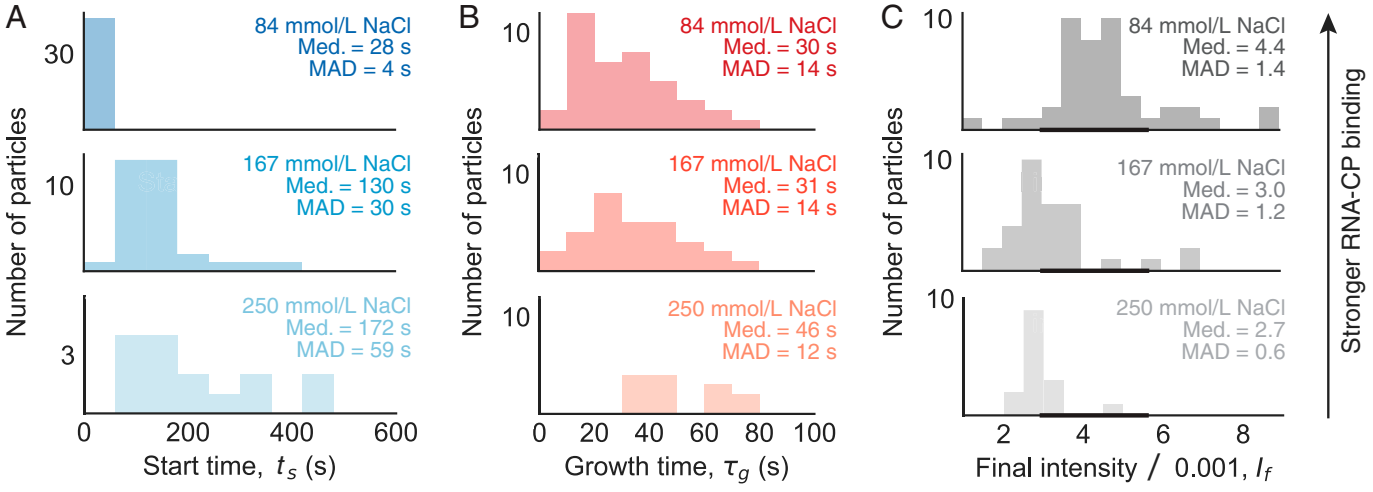


Fig. 5. RNA–protein interactions strongly affect nucleation but only weakly affect growth. Histograms of start time (t_s), growth time (τ_g), and final intensity (I_f) are plotted for three experiments with 0.135 μmol/L CP_2 and either 84 mmol/L, 167 mmol/L, or 250 mmol/L NaCl. (*A*) Start times t_s increase and broaden with decreasing RNA–protein interaction strength. (*B*) Growth times τ_g are less affected by RNA–protein interactions. (*C*) The final intensity I_f decreases with decreasing RNA–protein interaction strength. The bold portion of the *x*-axis in (*C*) shows the size range for full capsids. In (*A*–*C*), the median (Med.) and median absolute deviation (MAD) of each fit parameter are listed. Both the median and MAD are robust to outliers that arise from the few traces that are fit poorly by the piecewise function (see *SI Appendix*).

Downloaded from https://www.pnas.org by NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (NIST) on September 22, 2022 from IP address 129.6.219.63.

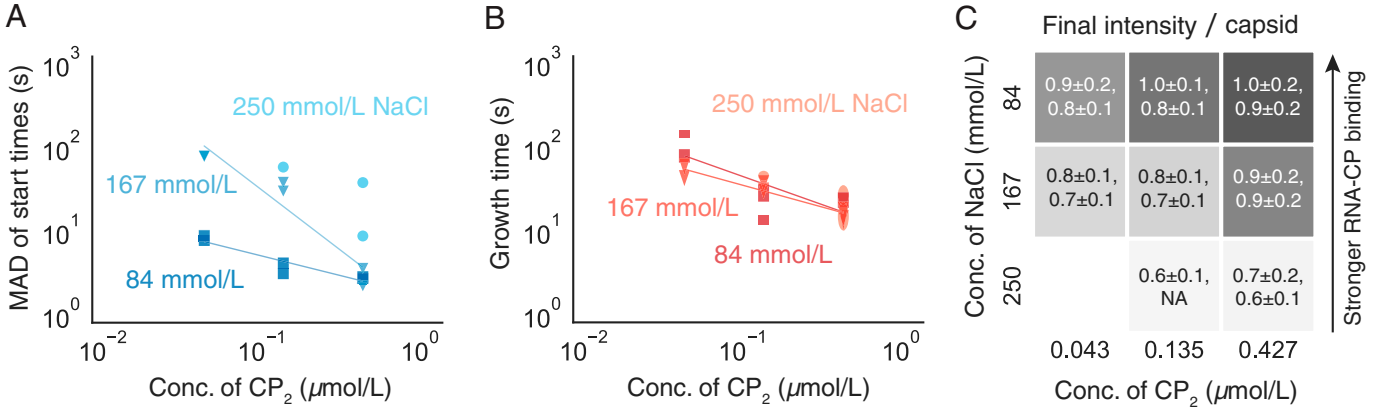


Fig. 6. Experiments at varying protein concentration suggest that nucleation occurs by a collective process involving multiple proteins, while growth occurs by a lower-order process. (*A*) A log–log plot of the spread of start times as a function of protein concentration and RNA–protein interaction strength (squares: 84 mmol/L NaCl; triangles: 167 mmol/L; circles: 250 mmol/L NaCl). Duplicate measurements are plotted for each set of conditions except for 250 mmol/L NaCl and 0.135 μmol/L CP_2, in which one of the duplicates did not yield assembly during the 600 s measurement time, and 250 mmol/L NaCl and 0.043 μmol/L CP_2, which was not tested. Lines have been fit to the data from the 84 mmol/L and 167 mmol/L NaCl experiments to show how the dependence on protein concentration becomes steeper at higher NaCl (no line is fitted to the 250 mmol/L data because only two protein concentrations were tested at that NaCl concentration). (*B*) Log–log plot of the median growth times shows a concentration dependence that is roughly the same for all ionic strengths (rectangles: 84 mmol/L NaCl; triangles: 167 mmol/L; ovals: 250 mmol/L NaCl). The height of the marker shows the MAD of each measurement. (*C*) Heat map of the median final intensity ± the MAD, normalized by the intensity of a full capsid, of each duplicate measurement (duplicates are separated by commas). The darkness of the gray color represents the magnitude of the average of the two median intensities. While the intensities tend to increase with increasing protein and decreasing NaCl, many of the assembled particles do not reach the size range of a full capsid (0.75–1.25), indicating that they are missing proteins. These smaller particles might have partial capsids that fail to completely close around the tethered RNA strand.

virus. The number and size of these aggregates decrease with decreasing RNA–protein interaction strength.

Discussion

We discuss our results for the self-assembly kinetics of individual BMV particles in comparison with previous results using the same technique to study bacteriophage MS2 (25), focusing first on their similarities. By highlighting similarities in the assembly kinetics, we aim to identify common features of the assembly process that might be found in other RNA viruses.

The assembly traces for BMV are strikingly similar to those previously reported for MS2 (Fig. 7) when RNA–protein interactions in BMV have been sufficiently weakened by high ionic strength (167 mmol/L or 250 mmol/L NaCl). For both viruses, we observe 1) broad distributions of start times that narrow with increasing protein concentration, consistent with a nucleation step; 2) growth times that decrease with increasing protein concentration but decrease less rapidly than the start times, consistent with growth involving a lower-order process; and 3) increases in the fraction of overgrown particles with increasing protein concentration, with overgrown particles consisting of aggregates of partially formed capsids. We note that the concentration of coat protein used in the assembly of MS2 is roughly a factor of 10 higher than that used with BMV, reflecting the stronger overall binding affinity of the BMV coat protein for RNA.

Another interesting similarity between BMV and MS2 assembly is that the apparent critical nucleus size is small—a few dimers—in both cases. We estimate this nucleus size from the amplitude of the fluctuations prior to the start time, as discussed in *SI Appendix*. In nearly all traces under conditions of moderate to weak RNA–protein interactions (167 mmol/L or 250 mmol/L NaCl), these fluctuations are comparable to the noise level, which corresponds to the intensity of a few CP_2. We use the term "apparent critical nucleus size" because averaging reduces our ability to observe large subcritical fluctuations. Nonetheless, this small apparent critical nucleus is comparable to that observed for MS2, and for both viruses we find that growth proceeds monotonically for nearly all traces, except for some small dips in the traces that are consistent with mechanical drift in the microscope (Fig. 7 *A* and *B*).

These results raise the question of how either BMV or MS2 is able to grow so rapidly, in a way that ensures that the capsid has the correct curvature, from such an apparently small nucleus. Nucleation by itself cannot solve the Levinthal paradox of virus assembly because the nucleus is too small to dictate the formation of a $T = 3$ structure during growth. Yet following nucleation, both BMV and MS2 display essentially monotonic growth to the size of a full capsid, with little if any disassembly or backtracking along the way. Furthermore, the curvature of the assembly products observed by TEM is similar to that of wild-type capsids even when the products are malformed aggregates of partially assembled capsids. How, then, do the coat proteins add to a growing nucleus, forming both the pentamers and hexamers needed for the proper curvature, with relatively few stalling events and without needing to detach from the particle?

A recent hypothesis is that specific subsequences on the RNA mediate the growth process to produce the correct curvature and structure. These subsequences are sometimes called packaging signals—a term that has come to refer to any elements of the RNA structure that have a high local affinity for coat protein. A model for MS2 assembly has been proposed in which upward of 60 strategically positioned packaging signals (18, 47) guide the assembly of the coat proteins. While not providing direct evidence in support of this model, high-resolution cryoelectron microscopy studies of MS2 and Qβ particles show approximately one dozen specific interactions between the RNA genome and its surrounding protein capsid (31, 32, 48–50), suggesting that packaging signals play a role in ordering the genome. But BMV is different. The absence of such order in the structure of the packaged BMV genome (51), and the lack of evidence of strategically positioned packaging signals in BMV RNA, weigh against there being a general paradigm that depends on a special distribution of specific RNA–protein interactions.

It is therefore interesting that we observe similar assembly kinetics for BMV and MS2. Our measurements suggest that

Downloaded from https://www.pnas.org by NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (NIST) on September 22, 2022 from IP address 129.6.219.63.

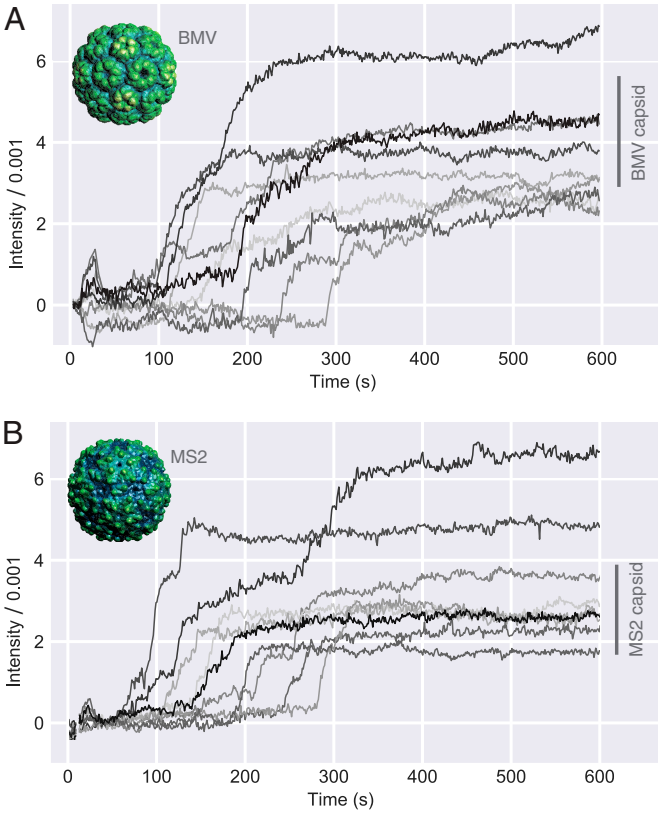


Fig. 7. Comparison of the assembly traces for BMV and MS2 reveals similar nucleation-and-growth kinetics. (*A*) BMV assembly traces for 0.043 μmol/L protein and 167 mmol/L NaCl. (*B*) For comparison, we show traces recorded for bacteriophage MS2 assembly under conditions of 2 μmol/L protein and 84 mmol/L NaCl (data from reference [25]). Both sets of traces reveal a broad distribution of start times followed by rapid and monotonic increases in intensity. In both sets of traces, most traces plateau to a final value that is roughly consistent with a full capsid, with several traces plateauing to larger values. The bars to the right of the traces show the range of intensities that are consistent with a full capsid. These ranges are different for MS2 and BMV, reflecting the different molecular masses of their coat protein.

whatever role packaging signals play in the self-assembly of these two viruses, their effect appears to be at most quantitative, with the qualitative features of the assembly process determined by general nonspecific interactions between the assembling proteins and the RNA. Determining the quantitative effects of packaging signals will ultimately require kinetic experiments in which the RNA sequence and structure are controlled (52). However, our experiments suggest that one should look to protein–protein interactions and intraprotein forces, rather than RNA–protein interactions, to understand how the capsid grows after nucleation. We find that the median growth time does not vary with NaCl concentration (Fig. 5*B*), indicating that RNA–protein interactions are not involved in the rate-limiting step of the growth phase of assembly and that either diffusion, protein–protein interactions, or intraprotein forces must therefore direct the rate-limiting step.

We glean additional information about the rate-limiting step from how the growth time depends on the protein concentration. We expect a scaling exponent equal to one if diffusion is rate limiting. By contrast, an exponent smaller than one might reflect a rate-limiting step that requires a change in conformation—either within the incoming subunit (53, 54) or between other subunits in the growing particle (55). At 84 mmol/L NaCl, the scaling exponent (0.7 ± 0.3; Fig. 6*B*) is consistent with diffusion-limited growth. At 167 mmol/L NaCl, the exponent (0.5 ± 0.1; Fig. 6*B*) leaves open the possibility of a more complex growth mechanism. Additional experiments over a broader range of concentrations could help pin down the precise relationship between growth time and protein concentration, which could help distinguish the operative mechanism.

Thus far we have discussed the similarities between BMV and MS2 assembly, but there are differences that arise when RNA–protein interactions are strong. At 84 mmol/L NaCl and 0.043 μmol CP_2, experiments on BMV assembly show a small spread in start times and relatively large growth times (see upper-left plot in Fig. 3). With MS2, we found that when the spread in start times is smaller than the average growth time, many oversized structures form. Our interpretation is that multiple nucleation events occur on the same RNA strand before the first nucleus can grow and sequester the RNA. With BMV, we observe similar overgrown structures but only at a higher protein concentration (Fig. 2*D*). We do not observe a significant number of overgrown structures at a low protein concentration. One possible explanation is that assembly is diffusion-limited when RNA–protein interactions are strong, occurring with no nucleation barrier, and that the small observed variation between traces is due to measurement noise. Alternatively, it is possible that the assembly is nucleated but the nucleation time is too small to resolve. In *SI Appendix*, we perform simulations to determine whether our measurements are consistent with any of the above scenarios—a diffusion-limited pathway subject to measurement noise or a nucleated pathway with a small nucleation time. These simulations suggest that our measurements are consistent with either of these pathways. Thus, additional experiments that can resolve smaller spreads in start times are needed to determine the assembly pathway when RNA–protein interactions are strong.

There also remains the question of if and when an ordered capsid arises from the RNA–protein complex. Interferometric scattering measurements do not address this question because the intensity depends primarily on the number of RNA-bound proteins and weakly, if at all, on their structure (27). There are at least two possible pathways leading to traces that do not show a clear separation between nucleation and growth. One is a barrier-less formation and growth of partially to completely ordered capsids, and the other is a barrier-less accretion of disordered proteins, followed by the onset of order among the bound proteins. These pathways cannot be distinguished in iSCAT because the measurement is blind to the onset of order. The same is true even when there is a clear separation between nucleation and growth processes. Here we observe a barrier to protein accretion, but it is possible that the proteins are not ordered. An analogous situation occurs in a bulk phase transition when gas-to-solid condensation takes place through a liquid-state intermediate. Thus, while our measurements can definitively resolve the presence of a barrier to protein accretion, future structural measurements are needed to distinguish whether the barrier involves ordered or disordered proteins.

Conclusions and Future Directions

The iSCAT experiments, with their high temporal resolution and ability to resolve the kinetics of assembly of individual viral capsids, offer the most detailed view to date of virus self-assembly pathways. Although iSCAT cannot reveal at what stage the capsid becomes ordered, much information can be gleaned by combining the technique with electron microscopy and bulk assays, as we have shown.

Downloaded from https://www.pnas.org by NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (NIST) on September 22, 2022 from IP address 129.6.219.63.

These experiments have shown that when RNA–protein binding is weak, the assembly kinetics of BMV are similar to those of MS2, involving a nucleation phase in which small numbers of proteins bind to the RNA, followed by a monotonic growth phase in which a capsid-worth of proteins steadily accrues. When RNA–protein binding is strong, we observe no clear barrier to protein accretion and hence no separation between nucleation and growth. Because our method may obscure nucleation events that occur on fast timescales, future studies are needed to determine whether for strong RNA–protein interactions there is a qualitatively different assembly pathway involving a saturated—"en masse" (15, 56)—adsorption of proteins on the RNA.

When the assembly is nucleated, we have shown that the time required to form a nucleus depends strongly on the strength of RNA-protein interactions, whereas the time needed to accrue a capsid-worth of proteins does not, suggesting that the RNA plays a more central role in the nucleation phase than in the growth phase. These results are consistent with protein accretion being a heterogeneous process in which subcritical protein clusters form on—and are stabilized by—the RNA. This role of the RNA degree of freedom—involving both RNA-protein binding and the conformation of the RNA itself (57)—compromises the notion of the nucleus being a fixed arrangement of coat proteins, such as a hexamer (58) or a pentamer (59) of dimers, although the uncertainties in our measurements cannot rule out this possibility. Future structural experiments, or kinetic experiments that address a broader range of timescales, may clarify whether the RNA enables different nucleus structures in different conditions.

Our experiments also reveal important features of the growth process. The shapes of the iSCAT traces show that growth can take place rapidly and without significant errors starting from an apparently small critical nucleus of only a few proteins. Furthermore, the weak dependence of growth times on NaCl concentration suggests that RNA–protein interactions do not control the kinetics of the growth phase and that protein–protein interactions and intraprotein forces may therefore direct the local curvature of the capsid and the emergence of the $T = 3$ structure. These results are compatible with simulations and theory showing that the growth of a $T = 3$ structure is driven by minimization of the elastic energy of the capsid (60, 61), which in general is related to the stretching and bending of coat-protein dimers as well as of the bonds between them. The elastic-energy hypothesis could explain why BMV and MS2 show similar assembly pathways.

To test this hypothesis, future experiments and analysis might focus on the growth process rather than nucleation. Recently, BMV and MS2 capsids of different sizes and symmetries have been observed (62–66). Experiments that determine how such malformed or overgrown capsids form would shed light on the interactions that control growth. Another useful next step is to develop coarse-grained computer simulations of the growth process (67–69). With such models, it would be possible to directly compare ensembles of simulated kinetic traces to the ensembles of traces measured with iSCAT. Agreement between these simulations and our data would point the way toward a detailed mechanism of the capsid formation process.

The commonalities in the in vitro assembly pathways of BMV and MS2 are remarkable in light of the vast phylogenetic distance between these viruses, their different RNA structures, and the differences in the specificity and strength of RNA–protein interactions. Such commonalities could reflect biological functions that are important for replication. For example, a free-energy barrier to nucleation could serve to delay RNA encapsulation until many copies of the viral proteins and genome have been produced. However, we do not know if the in vitro pathways we measure are operative in vivo, and there is reason to suspect that at least for MS2, there may be differences between the in vitro and in vivo pathways. The in vitro assemblies of MS2 are reconstituted from viral RNA and coat protein alone, whereas the wild-type MS2 particles assembled in vivo contain a single copy of a different gene product—the 44 kDa maturation protein—that replaces a coat protein dimer and breaks icosahedral symmetry.

Nonetheless, the existence of two different icosahedral, $T = 3$ capsids that assemble in vitro in similar ways is intriguing from both a physical and evolutionary perspective. From a physical perspective, this result suggests that the assembly of $T = 3$ viruses might be understood through a general physical theory. From an evolutionary perspective, it highlights the question of how icosahedral viruses with quasi-equivalent ($T > 1$) capsid subunits evolved. In MS2, for example, a single point mutation in the coat protein changes the structure of the capsid from $T = 3$ to $T = 1$ (66). Because capsids are self-assembled, such mutations can in principle also change the assembly pathway. Future studies might use interferometric scattering to explore how mutations affect the assembly of virus-like particles from different virus families, with the aim of discovering whether there are conserved interactions that promote robust assembly of the $T = 3$ structure.

Materials and Methods

Buffers used. Disassembly buffer: 50 mmol/L Tris·HCl, pH 7.5; 500 mmol/L $CaCl_2$; 1 mmol/L EDTA; 1 mmol/L DTT; 0.5 mmol/L PMSF. Protein storage buffer: 20 mmol/L Tris·HCl, pH 7.2; 1 mol/L NaCl; 1 mmol/L EDTA; 1 mmol/L DTT; 1 mmol/L PMSF. TAE buffer with NaCl: 40 mmol/L Tris·HCl, pH 8.3; 20 mmol/L acetic acid; 1 mmol/L EDTA; and 84 mmol/L, 167 mmol/L, or 250 mmol/L NaCl. Assembly buffer: 42 mmol/L 2-(*N*-morpholino)ethanesulfonic acid (MES), pH 6; 84 mmol/L, 167 mmol/L, or 250 mmol/L NaCl; 8.4 mmol/L $MgCl_2$; 3 mmol/L acetic acid. For DSF measurements, we also prepared assembly buffer at pH 7 by replacing MES with sodium phosphate.

Synthesis of BMV RNA1. BMV RNA1 (3,234 nucleotides) was made by in vitro transcription of the DNA plasmid pT7B1, linearized with BamHI (New England Biolabs),* with a T7 polymerase transcription system (Thermo Fisher) and purified with an RNEasy Mini Kit (Qiagen, DEU), both following the manufacturers' specifications.

BMV coat protein purification. BMV was purified from infected barley leaves (*Hordeum vulgare*) (70), and coat protein (189 aa, MW = 20,255 Da) was purified as described previously (71). We disassembled purified BMV particles by dialyzing against disassembly buffer at 4 °C overnight. We pelleted the RNA and isolated the coat protein by ultracentrifugation at 90,000 rotations per minute for 100 min at 4 °C in a Beckman TLA110 rotor. Coat protein was extracted from the supernatant and immediately dialyzed against the protein storage buffer. We used ultraviolet (UV) spectrophotometry to assess protein concentration and purity; only protein solutions with 260/280 ratios less than 0.6 were used for assembly. Protein was frozen in liquid nitrogen and stored at −80 °C until ready to use, at which point it was defrosted on ice and stored at 4 °C for up to 2 wk.

iSCAT mass photometry measurements. A solution of 3.5 μg/mL BMV coat protein in assembly buffer with either 84 mmol/L, 167 mmol/L, or 250 mmol/L NaCl was let to sit for 1 h at room temperature. We then diluted the protein solution two-fold by adding it to a drop of the corresponding assembly buffer sitting on top of an untreated glass coverslip. Proteins from the drop bound

*Certain commercial materials and equipment are identified in order to adequately specify the experimental procedure. Such identification does not imply recommendation by the National Institute of Standards and Technology.

Downloaded from https://www.pnas.org by NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (NIST) on September 22, 2022 from IP address 129.6.219.63.

nonspecifically to the glass surface and were imaged with a Refeyn TwoMP iSCAT microscope (Refeyn Ltd.). The change in intensity associated with each binding event was measured. We performed each experiment in triplicate at each buffer condition. To infer the molecular mass of a binding particle, and thus the number of 20 kDa coat-protein monomers within it, we compare its measured intensity to a calibration curve consisting of intensity measurements of three mass standards: 66 kDa BSA (Thermo Fisher) and 272 kDa trimers and 545 kDa hexamers of jack bean urease (MilliporeSigma). The resulting size distributions were normalized such that the area under each curve was equal to one. Each distribution reflected > 15,000 binding events.

Nitrocellulose-binding assay. Fluorescently labeled 20U RNA at a concentration of 4 nmol/L was mixed with 0.043 μmol/L BMV CP_2 in TAE buffer with either 84 mmol/L, 167 mmol/L, or 250 mmol/L NaCl and left for 30 min at room temperature. The RNA was labeled at its 5′-end with an AlexaFluor647 dye (Integrated DNA Technologies). Next, 250 μL aliquots of each RNA–protein mixture were passed through a nitrocellulose membrane (0.45 μm pore size; Thermo Fisher) that was presoaked in TAE buffer with 250 mmol/L NaCl using a 96-well dot-blot apparatus (Biorad) under weak vacuum. Then we washed the membrane by passing through an additional 500 μL corresponding assembly buffer. We quantified the amount of membrane-bound RNA at each NaCl concentration by measuring the fluorescence emission intensity of the AlexaFluor647 dye using a fluorescence scanner. The amount of protein bound to the membrane was constant for all NaCl concentrations, as determined by staining with Ponceau S solution (MilliporeSigma). Because protein is retained by the membrane but free RNA is not, we took the intensity of membrane-bound RNA to be proportional to the yield of RNA–protein binding.

DSF. Wild-type BMV was dialyzed against assembly buffer at pH 6 or pH 7 overnight at 4 °C. Aliquots of wild-type BMV at a final concentration of 0.2 mg/mL, 2.5× SYPRO orange fluorescent dye (Molecular Probes), and a final NaCl concentration of 85 mmol/L, 167 mmol/L, or 250 mmol/L NaCl were prepared. DSF was performed in triplicate in a 96-well plate CFX Connect quantitative PCR machine (Biorad). All samples were heated from 25 °C to 95 °C, in 1 °C increments with a 1-min stabilization period at each temperature before the sample was measured. Excitation/emission wavelengths of 470/550 nm were used to detect the fluorescence emission of SYPRO orange binding to hydrophobic regions of coat protein exposed upon capsid disassembly. The apparent melting temperature was defined as the temperature at which the derivative of the fluorescence emission signal ($-dI/dT$) is maximal.

Negative-stain electron microscopy. Negative-stain electron microscopy was used to image the protein structures that assemble around RNA in solution. We carried out the assembly reaction in assembly buffer by mixing 860 nmol/L BMV CP_2 with 7.5 nmol/L BMV RNA1 and incubating at room temperature for 10 min. We deposited 6 μL assembly reaction on glow-discharged carbon-coated copper (200-mesh) PELCO Pinpointer grids (Ted Pella). After 1 min, the grids were blotted with Whatman filter paper and then stained with 6 μL 2% uranyl acetate for 1 min followed by complete stain removal and storage in a desiccator overnight. Micrographs were acquired using a Tecnai G2 TF20 High-Resolution electron microscope (FEI) with an accelerating voltage of 200 kV. Images were collected at 3 μm to 4 μm underfocus with a TIETZ F415MP 16-megapixel CCD camera (4,000 by 4,000 pixels, pixel size 15 μm).

iSCAT assembly kinetics measurements. Time-resolved iSCAT measurements of capsid assembly were performed with a custom-built iSCAT microscope, as described in detail in reference (25) and the accompanying *SI Appendix* to reference (25). In brief, a spatially filtered, 450-nm diode laser was coupled into an oil–immersion objective to illuminate a small region of the coverslip with a collimated beam. Three-dimensional active stabilization was used to extend the measurement duration.

Coverslips were functionalized with a layer of PEG molecules, approximately 1% of which were functionalized with a 20-base DNA strand (5′-GGTT GGTTGGTTGGTTGGTT-3′), to which BMV RNA1 strands were tethered with a 60-base DNA linker strand (5′-CCGTGGTCGACAAGGGATTGAACCTCGTTCCGTGGTC TACAACCAACCAACCAACCAACC-3′). The 40 bases at the 5′-end of the linker strand bound the 5′-end of BMV RNA1, and the remaining 20 bases bound the surface-tethered DNA. Previously, we showed that ordered BMV capsids can assemble around tethered RNA, with the DNA linkage poking out of the capsid through a defect too small to visualize by TEM (26).

Assembly kinetics experiments were performed at room temperature in assembly buffer with 84 mmol/L, 167 mmol/L, and 250 mmol/L NaCl and protein concentrations of 0.043 μmol/L, 0.135 μmol/L, and 0.427 μmol/L BMV CP_2. The diffraction-limited spot size was ∼150 nm, much larger than either a BMV particle or BMV RNA1 (35). For each assembling particle, we manually located the center and measured the mean intensity in a circle of radius 1 pixel about that center. We estimated the intensity range of full BMV capsids by scaling the measured intensity range for MS2 capsids (25) by the relative mass of BMV and MS2 capsids.

Data, Materials, and Software Availability. All iSCAT intensity data are included as *SI Appendix*. The raw movie files from which we extracted the intensity data will be made available upon request.

ACKNOWLEDGMENTS. We thank Ben Rogers and Abigail Chapman for many helpful discussions. Research reported in this publication was supported by the National Institute of General Medical Sciences of the NIH (award number R00GM127751 to R.F.G.) and the NSF Molecular and Cellular Biosciences Division (awards MCB 1716925 and 2103700 to W.M.G.). This research was partially supported by NSF through the Harvard University Materials Research Science and Engineering Center under NSF grants DMR-1420570 and DMR-2011754. A.M.G. was partially supported by an NSF Graduate Research Fellowship under grant DGE-1144152. We acknowledge the California Metabolic Research Foundation for its support of biochemical research at San Diego State University and the Bauer Core Facility at Harvard University for shared experimental facilities used in this study. Portions of the paper were developed from the thesis of C.R.T.

Author affiliations: [a]Harvard John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138; [b]Department of Chemistry and Biochemistry, San Diego State University, San Diego, CA 92182; [c]Viral Information Institute, San Diego State University, San Diego, CA 92182; [d]Sensor Science Division, National Institute of Standards and Technology, Gaithersburg, MD 20899; [e]Department of Chemistry and Biochemistry, University of California Los Angeles, CA 90095; [f]Department of Chemistry, Colorado School of Mines, Golden, CO 80401; [g]Molecular Biology Institute, University of California Los Angeles, CA 90095; [h]California Nanosystems Institute, University of California Los Angeles, CA 90095; and [i]Department of Physics, Harvard University, Cambridge, MA 02138

1. F. H. C. Crick, J. D. Watson, Structure of small viruses. *Nature* **177**, 473–475 (1956).
2. D. L. Caspar, A. Klug, Physical principles in the construction of regular viruses. *Cold Spring Harb. Symp. Quant. Biol.* **27**, 1–24 (1962).
3. S. J. Stray *et al.*, A heteroaryldihydropyrimidine activates and can misdirect hepatitis B virus capsid assembly. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 8138–8143 (2005).
4. D. S. Peabody *et al.*, Immunogenic display of diverse peptides on virus-like particles of RNA phage MS2. *J. Mol. Biol.* **380**, 252–263 (2008).
5. G. L. Butterfield *et al.*, Evolution of a designed protein assembly encapsulating its own RNA genome. *Nature* **552**, 415–420 (2017).
6. M. G. Rossmann, J. E. Johnson, Icosahedral RNA virus structure. *Annu. Rev. Biochem.* **58**, 533–573 (1989).
7. R. Twarock, A. Luque, Structural puzzles in virology solved with an overarching icosahedral design principle. *Nat. Commun.* **10**, 4414 (2019).
8. A. Zlotnick, J. M. Johnson, P. W. Wingfield, S. J. Stahl, D. Endres, A theoretical model successfully identifies features of hepatitis B virus capsid assembly. *Biochemistry* **38**, 14644–14652 (1999).
9. S. Mukhopadhyay, P. R. Chipman, E. M. Hong, R. J. Kuhn, M. G. Rossmann, In vitro-assembled alphavirus core-like particles maintain a structure similar to that of nucleocapsid cores in mature virus. *J. Virol.* **76**, 11128–11132 (2002).
10. M. Comas-Garcia *et al.*, Efficient support of virus-like particle assembly by the HIV-1 packaging signal. *eLife* **7**, e38438 (2018).
11. Y. I. Wolf *et al.*, Origins and evolution of the global RNA virome. *mBio* **9**, e0239-18 (2018).
12. J. E. Johnson, J. A. Speir, Quasi-equivalent viruses: A paradigm for protein assemblies. *J. Mol. Biol.* **269**, 665–675 (1997).
13. J. B. Bancroft, E. Hiebert, Formation of an infectious nucleoprotein from protein and nucleic acid isolated from a small spherical virus. *Virology* **32**, 354–356 (1967).
14. J. B. Bancroft, G. J. Hills, R. Markham, A study of the self-assembly process in a small spherical virus. Formation of organized structures from protein subunits in vitro. *Virology* **31**, 354–379 (1967).
15. E. Hiebert, J. B. Bancroft, C. E. Bracker, The assembly in vitro of some small spherical viruses, hybrid viruses, and other nucleoproteins. *Virology* **34**, 492–508 (1968).
16. T. Sugiyama, R. R. Hebert, K. A. Hartman, Ribonucleoprotein complexes formed between bacteriophage MS2 RNA and MS2 protein in vitro. *J. Mol. Biol.* **25**, 455–463 (1967).
17. E. C. Dykeman, P. G. Stockley, 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).

Downloaded from https://www.pnas.org by NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (NIST) on September 22, 2022 from IP address 129.6.219.63.

18. R. Twarock, R. J. Bingham, E. C. Dykeman, P. G. Stockley, A modelling paradigm for RNA virus assembly. *Curr. Opin. Virol.* **31**, 74–81 (2018).
19. B. Dragnea, Watching a virus grow. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 22420–22422 (2019).
20. A. Borodavka, R. Tuma, P. G. Stockley, Evidence that viral RNAs have evolved for efficient, two-stage packaging. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 15769–15774 (2012).
21. A. Borodavka, R. Tuma, P. G. Stockley, A two-stage mechanism of viral RNA compaction revealed by single molecule fluorescence. *RNA Biol.* **10**, 481–489 (2013).
22. R. F. Garmann, M. Comas-Garcia, A. Gopal, C. M. Knobler, W. M. Gelbart, The assembly pathway of an icosahedral single-stranded RNA virus depends on the strength of inter-subunit attractions. *J. Mol. Biol.* **426**, 1050–1060 (2014).
23. M. Chevreuil *et al.*, Nonequilibrium self-assembly dynamics of icosahedral viral capsids packaging genome or polyelectrolyte. *Nat. Commun.* **9**, 3071 (2018).
24. J. Ortega-Arroyo, P. Kukura, Interferometric scattering microscopy (iSCAT): New frontiers in ultrafast and ultrasensitive optical microscopy. *Phys. Chem. Chem. Phys.* **14**, 15625–15636 (2012).
25. R. F. Garmann, A. M. Goldfain, 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).
26. R. F. Garmann *et al.*, A simple RNA-DNA scaffold templates the assembly of monofunctional virus-like particles. *J. Am. Chem. Soc.* **137**, 7584–7587 (2015).
27. G. Young *et al.*, Quantitative mass imaging of single biological macromolecules. *Science* **360**, 423–427 (2018).
28. J. Carey, V. Cameron, P. L. de Haseth, O. C. Uhlenbeck, Sequence-specific interaction of R17 coat protein with its ribonucleic acid binding site. *Biochemistry* **22**, 2601–2610 (1983).
29. D. Beckett, H. N. Wu, O. C. Uhlenbeck, Roles of operator and non-operator RNA sequences in bacteriophage R17 capsid assembly. *J. Mol. Biol.* **204**, 939–947 (1988).
30. G. G. Pickett, D. S. Peabody, Encapsidation of heterologous RNAs by bacteriophage MS2 coat protein. *Nucleic Acids Res.* **21**, 4621–4626 (1993).
31. R. I. Koning *et al.*, Asymmetric cryo-EM reconstruction of phage MS2 reveals genome structure in situ. *Nat. Commun.* **7**, 12524 (2016).
32. X. Dai *et al.*, In situ structures of the genome and genome-delivery apparatus in a single-stranded RNA virus. *Nature* **541**, 112–116 (2017).
33. P. Ni *et al.*, 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).
34. Y. G. Choi, T. W. Dreher, A. L. N. Rao, tRNA elements mediate the assembly of an icosahedral RNA virus. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 655–660 (2002).
35. R. F. Garmann, M. Comas-Garcia, C. M. Knobler, W. M. Gelbart, Physical principles in the self-assembly of a simple spherical virus. *Acc. Chem. Res.* **49**, 48–55 (2016).
36. R. F. Garmann *et al.*, Role of electrostatics in the assembly pathway of a single-stranded RNA virus. *J. Virol.* **88**, 10472–10479 (2014).
37. P. Pfeiffer, L. Hirth, Aggregation states of brome mosaic virus protein. *Virology* **61**, 160–167 (1974).
38. A. Strugała *et al.*, Virus-like particles produced using the brome mosaic virus recombinant capsid protein expressed in a bacterial system. *Int. J. Mol. Sci.* **22**, 3098 (2021).
39. M. Hema *et al.*, Effects of amino-acid substitutions in the brome mosaic virus capsid protein on RNA encapsidation. *Mol. Plant Microbe Interact.* **23**, 1433–1447 (2010).
40. V. Rayaprolu *et al.*, Fluorometric estimation of viral thermal stability. *Bio Protoc.* **4**, e1199 (2014).
41. F. Tama, C. L. Brooks III, The mechanism and pathway of pH induced swelling in cowpea chlorotic mottle virus. *J. Mol. Biol.* **318**, 733–747 (2002).
42. P. G. Stockley, Filter-binding assays. *Methods Mol. Biol.* **30**, 251–262 (1994).
43. M. F. Hagan, A theory for viral capsid assembly around electrostatic cores. *J. Chem. Phys.* **130**, 114902 (2009).
44. R. Zandi, P. van der Schoot, D. Reguera, W. Kegel, H. Reiss, Classical nucleation theory of virus capsids. *Biophys. J.* **90**, 1939–1948 (2006).
45. T. C. T. Michaels, A. J. Dear, T. P. J. Knowles, Stochastic calculus of protein filament formation under spatial confinement. *New J. Phys.* **20**, 055007 (2018).
46. A. Hensley, W. M. Jacobs, W. B. Rogers, Self-assembly of photonic crystals by controlling the nucleation and growth of DNA-coated colloids. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2114050118 (2022).
47. P. G. Stockley *et al.*, Packaging signals in single-stranded RNA viruses: Nature's alternative to a purely electrostatic assembly mechanism. *J. Biol. Phys.* **39**, 277–287 (2013).
48. K. V. Gorzelnik *et al.*, Asymmetric cryo-EM structure of the canonical allolevivirus Qβ reveals a single maturation protein and the genomic ssRNA in situ. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 11519–11524 (2016).
49. Z. Cui *et al.*, Structures of Qβ virions, virus-like particles, and the Qβ-MurA complex reveal internal coat proteins and the mechanism of host lysis. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 11697–11702 (2017).
50. J.-Y. Chang, K. V. Gorzelnik, J. Thongchol, J. Zhang, Structural assembly of Qβ virion and its diverse forms of virus-like particles. *Viruses* **14**, 225 (2022).
51. C. Beren *et al.*, Genome organization and interaction with capsid protein in a multipartite RNA virus. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 10673–10680 (2020).
52. Y. Song *et al.*, Limits of variation, specific infectivity, and genome packaging of massively recoded poliovirus genomes. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E8731–E8740 (2017).
53. S. Asakura, G. Eguchi, T. Iino, Reconstitution of bacterial flagella in vitro. *J. Mol. Biol.* **10**, 42–56 (1964).
54. S. Asakura, A kinetic study of in vitro polymerization of flagellin. *J. Mol. Biol.* **35**, 237–239 (1968).
55. D. L. Caspar, Movement and self-control in protein assemblies. Quasi-equivalence revisited. *Biophys. J.* **32**, 103–138 (1980).
56. J. D. Perlmutter, M. R. Perkett, M. F. Hagan, Pathways for virus assembly around nucleic acids. *J. Mol. Biol.* **426**, 3148–3165 (2014).
57. S. W. Singaram, R. F. Garmann, C. M. Knobler, W. M. Gelbart, A. Ben-Shaul, Role of RNA branchedness in the competition for viral capsid proteins. *J. Phys. Chem. B* **119**, 13991–14002 (2015).
58. J. A. Speir, S. Munshi, G. Wang, T. S. Baker, J. E. Johnson, Structures of the native and swollen forms of cowpea chlorotic mottle virus determined by X-ray crystallography and cryo-electron microscopy. *Structure* **3**, 63–78 (1995).
59. A. Zlotnick, R. Aldrich, J. M. Johnson, P. Ceres, M. J. Young, Mechanism of capsid assembly for an icosahedral plant virus. *Virology* **277**, 450–456 (2000).
60. S. Li, P. Roy, A. Travesset, R. Zandi, Why large icosahedral viruses need scaffolding proteins. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 10971–10976 (2018).
61. S. Panahandeh *et al.*, How a virus circumvents energy barriers to form symmetric shells. *ACS Nano* **14**, 3170–3180 (2020).
62. K. Bond, I. B. Tsvetkova, J. C.-Y. Wang, M. F. Jarrold, B. Dragnea, Virus assembly pathways: Straying away but not too far. *Small* **16**, e2004475 (2020).
63. N. de Martín Garrido *et al.*, Bacteriophage MS2 displays unreported capsid variability assembling T = 4 and mixed capsids. *Mol. Microbiol.* **113**, 143–152 (2020).
64. A. P. Biela, A. Naskalska, F. Fatehi, R. Twarock, J. G. Heddle, Programmable polymorphism of a virus-like particle. *Commun Mater* **3**, 7 (2022).
65. M. Ruszkowski *et al.*, Cryo-EM reconstructions of BMV-derived virus-like particles reveal assembly defects in the icosahedral lattice structure. *Nanoscale* **14**, 3224–3233 (2022).
66. M. A. Asensio *et al.*, A selection for assembly reveals that a single amino acid mutant of the bacteriophage MS2 coat protein forms a smaller virus-like particle. *Nano Lett.* **16**, 5944–5950 (2016).
67. M. F. Hagan, "Modeling viral capsid assembly" in *Advances in Chemical Physics*, S. A. Rice, A. R. Dinner, Eds. (John Wiley & Sons, Inc., 2014), **vol. 155**, pp. 1–68.
68. J. D. Perlmutter, M. F. Hagan, Mechanisms of virus assembly. *Annu. Rev. Phys. Chem.* **66**, 217–239 (2015).
69. M. F. Hagan, R. Zandi, Recent advances in coarse-grained modeling of virus assembly. *Curr. Opin. Virol.* **18**, 36–43 (2016).
70. J. B. Bancroft, The self-assembly of spherical plant viruses. *Adv. Virus Res.* **16**, 99–134 (1970).
71. P. Annamalai, A. L. N. Rao, Dispensability of 3′ tRNA-like sequence for packaging cowpea chlorotic mottle virus genomic RNAs. *Virology* **332**, 650–658 (2005).
72. C. R. Tanimoto, *"The physical and chemical parameters of in vitro packaging of heterologous RNA into virus-like particles,"* Thesis, University of California Los Angeles, CA (2022).

Downloaded from https://www.pnas.org by NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (NIST) on September 22, 2022 from IP address 129.6.219.63.