JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
Cite This: J. Am. Chem. Soc. XXXX, XXX, XXX−XXX
pubs.acs.org/JACS

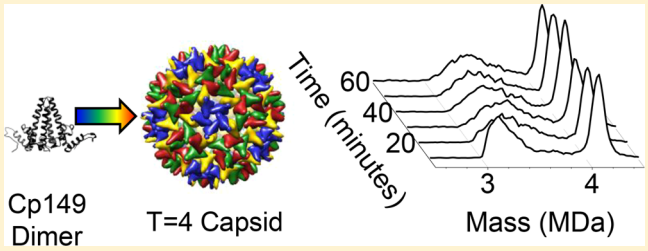
Multiple Pathways in Capsid Assembly

Corinne A. Lutomski,[†] Nicholas A. Lyktey,[†] Elizabeth E. Pierson,[†] Zhongchao Zhao,[‡] Adam Zlotnick,*[,‡] and Martin F. Jarrold*[,†]

[†]Department of Chemistry and [‡]Department of Molecular and Cellular Biochemistry, Indiana University, Bloomington, Indiana 47405, United States

Supporting Information

ABSTRACT: For a three-dimensional structure to spontaneously self-assemble from many identical components, the steps on the pathway must be kinetically accessible. Many virus capsids are icosahedral and assembled from hundreds of identical proteins, but how they navigate the assembly process is poorly understood. Capsid assembly is thought to involve stepwise addition of subunits to a growing capsid fragment. Coarse-grained models suggest that the reaction occurs on a downhill energy landscape, so intermediates are expected to be fleeting. In this work, charge detection mass spectrometry (CDMS) has been used to track assembly of the hepatitis B virus (HBV) capsid in real time. The icosahedral $T = 4$ capsid of HBV is assembled from 120 capsid protein dimers. Our results indicate that there are multiple pathways for assembly. Under conditions that favor a modest association energy there is no accumulation of large intermediates, which indicates that available pathways include ones on a downhill energy surface. Under higher salt conditions, where subunit interactions are strengthened, around half of the products of the initial assembly reaction have masses close to the $T = 4$ capsid and the other half are stalled intermediates which emerge abruptly at around 90 dimers, indicating a bifurcation in the ensemble of assembly paths. When incubated at room temperature, the 90-dimer intermediates accumulate dimers and gradually shift to higher mass and merge with the capsid peak. Though free subunits are present in solution, the stalled intermediates indicate the presence of a local minima on the energy landscape. Some intermediates may result from hole closure, where the growing capsid distorts to close the hole due to the missing capsid proteins or from a species where subsequent additions are particularly labile.



INTRODUCTION

The design of molecular components that self-assemble into a specific three-dimensional structure is a challenging task because the geometry of the structure must be encoded into the intermolecular interactions of the components. With recent developments in synthetic biology it is now possible to design proteins that assemble into small cages.[1−8] However, the most exquisite examples of protein cages still come from nature. A virus capsid, the protein shell that surrounds the viral genome, is assembled from many, often hundreds, of identical proteins. For self-assembly to occur, the desired end point must be thermodynamically favored and kinetically accessible. In other words, there must be facile, low-energy pathways that lead to the desired structure.[9,10] While a lot is known about the structures of natural protein cages, much less is known about how they assemble and avoid undesirable outcomes.

A spherical cage structure can, in theory, be generated with any number of components.[11,12] However, spherical virus capsids are generally icosahedral and found in discrete sizes. The sizes are described by a triangulation number (T), where $60T$ is the number of proteins in the capsid.[13] Symmetry dictates that only certain T numbers are allowed; the smallest ones are 1, 3, 4, and 7, and they contain 60, 180, 240, and 420 capsid proteins, respectively.

Capsid assembly is thought to involve three steps: formation of a nucleus (a small capsid fragment), the sequential addition of capsid building blocks to the growing capsid, and finally closure, where the final building block is inserted to complete the icosahedron.[14−18] The building block can be a capsid protein, a dimer, or a larger capsid fragment. Coarse-grained models suggest that the energy landscape for subunit addition to the nucleus is downhill,[18] except perhaps for the completion step.[16] The energy landscape for self-assembly has been visualized as a "palm tree". Branches corresponding to higher energy local minima with low barriers form a steep potential energy gradient that effectively funnel the system into a well-defined global minimum.[9,10,12,19−21] During the assembly reaction, capsid building blocks and fully formed capsids are expected to be dominant over low abundance intermediates.[22] Capsid disassembly is a similarly complex process where up to ∼25% of subunits maybe removed, leaving a 90-dimer complex, until a catastrophic dissociation at a percolation threshold.[23]

In the work reported here we have used charge detection mass spectrometry (CDMS)[24,25] to track the assembly of the icosahedral $T = 4$ capsid of hepatitis B virus (HBV) in real time. Infectious HBV virions are formed through the assembly of the

Received: February 13, 2018

© XXXX American Chemical Society

DOI: 10.1021/jacs.8b01804
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

wild type 183-residue capsid protein (Cp183) nucleated by a nucleic acid−polymerase complex.[26,27] However, about 90% of virus-like particles found in vivo are empty, demonstrating efficient self-assembly and indicating the importance of empty particles in infections.[28] A truncated form of the capsid protein which lacks the C-terminal nucleic acid binding domain (Cp149) also assembles into icosahedral capsids[29] that are indistinguishable from authentic virus particles.[30,31] Capsids are believed to assemble by sequential addition of Cp149 dimers to a nucleus.[32] In vitro assembly can be triggered by raising the ionic strength, which strengthens dimer−dimer interactions. However, at high ionic strengths intermediates may become kinetically trapped. This can occur when a large number of nuclei are formed early in the assembly reaction and there is insufficient dimer present to complete them.

High-resolution mass spectrometry and ion mobility mass spectrometry have been used to investigate early intermediates in HBV capsid assembly as well as intact capsids.[33−35] Using charge detection mass spectrometry (CDMS), we previously identified kinetically trapped, high-mass intermediates in HBV assembly.[36] CDMS is a single-particle technique where the mass of each ion is directly determined from simultaneous measurements of its mass-to-charge ratio (m/z) and charge (z). It is particularly well-suited for monitoring the heterogeneous mixture of high-mass intermediates that may result from capsid assembly.

EXPERIMENTAL METHODS

Preparation of Capsid Protein Cp149. The truncated form of the capsid protein containing only the assembly domain (Cp149) was expressed in *E. coli* and purified as previously described.[37] A reassembly and dissociation step was included in the purification procedure to remove inactive protein. Assembly competent protein was dialyzed into 20 mM ammonium acetate at pH 7.5. Assembly was initiated by raising the ammonium acetate concentration to increase the ionic strength.

Time-Resolved Charge Detection Mass Spectrometry. In ion trap CDMS, single ions are trapped in a linear ion trap where they oscillate back and forth through a conducting cylinder. As an ion enters the cylinder it is detected by a charge sensitive amplifier. The resulting signal is analyzed to provide the m/z and charge, which are combined to provide the mass of each ion.

A detailed description of the home-built charge detection mass spectrometer used in this work can be found elsewhere.[38−42] Ions are generated by a chip-based nanoelectrospray source. They enter the instrument through a heated capillary and pass through three differentially pumped regions where they are separated from the ambient gas that flows through the capillary. The ions are then accelerated through a 100 V potential difference and focused into an ion beam. A narrow band of ion energies centered on 100 eV/z is selected by a dual hemispherical deflection energy analyzer (HDA) and focused into a modified cone trap that contains the charge detection cylinder. Potentials are applied to the end-cap electrodes to trap ions. In this work we used continuous or random trapping (where the trap is closed without knowing whether an ion is present). The trapping period was 100 ms.

As the trapped ion oscillates back and forth in the detection cylinder it induces a periodic signal which is amplified, digitized, and then analyzed in real time by a Fortran program using fast Fourier transforms. The m/z is obtained from the fundamental frequency, and the charge is obtained from the magnitudes of the fundamental and first harmonic. The calibration of the charge and m/z measurements has been discussed elsewhere.[42] Ions passing through the trap have a narrow distribution of kinetic energy per charge defined by the HDA, and therefore each ion's velocity is related to its m/z, i.e., $v \propto (m/z)^{-1/2}$. With random trapping the probability an ion is trapped is proportional to the time it spends in the trap. Consequently, slow (high m/z) ions are trapped more easily than fast (low m/z) ions. To correct for this difference, each ion is weighted by its $m/z^{-1/2}$.

For the time-resolved CDMS measurements, a solution of 10, 20, or 40 μM Cp149 dimer in 20 mM ammonium acetate is mixed with an equal volume of 400 or 1000 mM ammonium acetate solution to initiate assembly. After mixing, the reaction mixture is immediately loaded into the reservoir of the nanoelectrospray source, where the assembly reaction continues while the solution is electrosprayed. The ions detected by CDMS are time-stamped, and so knowing when the assembly reaction was initiated, it is possible to relate each ion to a particular reaction time. The frequency of single ion trapping events is relatively low, 3−4 ions per second at best, with random trapping and the 100 ms trapping period used here. Too few ions are detected in a single reaction to generate time-resolved spectra. To increase the number of ions, the results of many identical reactions are combined and sorted into time windows.

RESULTS

Time-resolved CDMS measurements were performed with a range of ammonium acetate and Cp149 dimer concentrations. Kinetic traps can be induced in reactions at relatively high protein concentrations or at relatively high association energy proportional to ionic strength. The measurements in this series of experiments were performed with Cp149 dimer concentrations (after mixing) of 5, 10, and 20 μM and ammonium acetate concentrations of 210 and 510 mM.

The results in each panel of Figure 1 are a compilation of the ions measured for 8−28 identical experiments. Ions were recorded continuously for at least 90 min after the assembly reaction was initiated. Some reaction mixtures were also sampled periodically over the course of several days. The x-axis in Figure 1 shows the mass parsed into 20 kDa bins, about half the mass of a dimer, the y-axis is the intensity normalized by total peak area including low mass ions (not shown in the plot for clarity), and the z-axis represents the reaction time binned into 10 min intervals. The earliest time interval (0−10 min) is at the front of each plot.

Figure 1A shows time-resolved CDMS spectra measured over the first 90 min of an assembly reaction with a dimer concentration of 5 μM and an ammonium acetate concentration of 210 mM. The most prominent assembly product in the CDMS spectra is the peak at around 4.1 MDa, which is attributed to an overgrown T = 4 capsid (i.e., a capsid with slightly more than the expected 120 dimers).[43] A small amount of intermediate of ≥90 dimers is also observed. HBV can form both 90-dimer T = 3 and 120-dimer T = 4 capsids in vivo and in vitro.[44,45] In vivo, around 5% of the capsids are T = 3.[46] For in vitro assembly with Cp149, the relative abundances are affected by the identity and ionic strength of the assembly buffer.[47,48] In ammonium acetate, little T = 3 capsid is formed under low salt conditions.[43] Therefore, we expect that most of this broad peak is due to intermediates.

If the assembly reaction is treated as an equilibrium, the capsid concentration is given by

$$[\text{capsid}] = K[\text{dimer}]^{120}$$

The large stoichiometry exponent leads to a pseudo-critical concentration for capsid formation. Below the pseudo-critical concentration the capsid concentration is very small; free dimer is the main species present. Above the pseudo-critical concentration, the amount of capsid increases approximately linearly as any dimer above the pseudo-critical concentration is mainly converted into capsid.[22] Under the conditions used here, the pseudo-critical concentration is 3.55 μM. Thus, the

DOI: 10.1021/jacs.8b01804
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

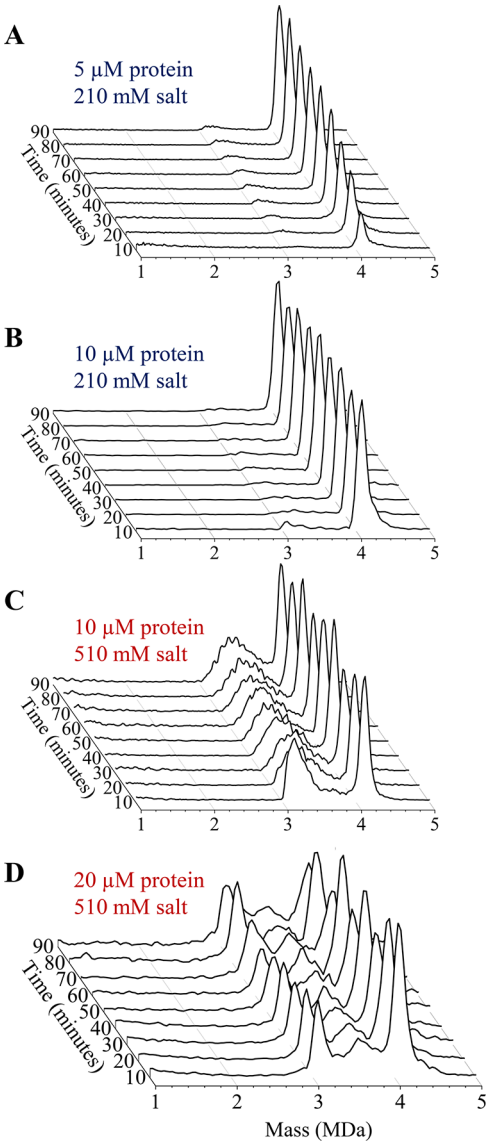


Figure 1. Time-resolved CDMS spectra showing the progression of capsid assembly over the first 90 min for a reaction mixture containing an initial dimer concentration of (A) 5 μM in 210 mM ammonium acetate, (B) 10 μM in 210 mM ammonium acetate, (C) 10 μM in 510 mM ammonium acetate, and (D) 20 μM in 510 mM ammonium acetate. The earliest time interval (0−10 min) is at the front. The spectra were generated using 20 kDa bins. Each spectrum is normalized by the sum of all ions measured for each time point.

dimer concentration used to obtain the results in Figure 1A (5 μM) is only slightly above the pseudo-critical concentration.

The low dimer concentration coupled with the low salt concentration used to obtain the results in Figure 1A leads to relatively slow assembly. The peak at 4.1 MDa can be seen to grow during the 90 min time window shown. We found that different dimer preparations often led to different rates of assembly. The results shown in Figure 1A are on the slow end for these conditions. A few other species are observed in the 1−5 MDa mass range shown in the figure. In addition to the broad peak at 3.2 MDa, there is a broad low-intensity feature at early times at around 1 MDa discussed below.

Figure 1B shows the result of increasing the dimer concentration to 10 μM under the same low salt conditions used for Figure 1A. With the higher dimer concentration, the initial assembly reaction is much faster and appears to be complete well within the first 10 min time interval. A small feature at around 3.1 MDa is present in the first time point and shifts to higher mass as the reaction proceeds.

Figure 1C shows the effect of increasing the salt concentration while maintaining the dimer concentration (10 μM) used in Figure 1B. Raising the salt concentration leads to a large increase in the abundances of species with masses between 3 and 4 MDa. According to light scattering measurements, the initial assembly reaction approaches equilibrium in under a minute for the conditions used here.[43] A little over half of the of the products have masses between 3 and 4 MDa while a little under half are in the overgrown capsid peak at around 4.1 MDa. At the first time point in Figure 1C there is a broad feature at around 3.2 MDa with a high mass tail and a sharper capsid peak at around 4.1 MDa. As the reaction proceeds, the broad peak at around 3.2 MDa becomes broader and moves to higher masses, progressing toward the capsid peak. Nonetheless, after 90 min, most of the large species with masses between 3 and 4 MDa are still present, and there has not been a significant increase in the intensity of the capsid peak. The variations in the intensity of the peak at around 4 MDa in Figure 1C are attributed to random fluctuations from the way the spectra are assembled from many individual experiments. The variations diminish as more experiments are summed.

Figure 1D shows time-resolved CDMS spectra measured over the first 90 min of an assembly reaction initiated with a Cp149 dimer concentration of 20 μM and an ammonium acetate concentration of 510 mM. Doubling the dimer concentration under high salt conditions did not lead to a significant change in the relative abundance of the capsid peak and dimer (not shown). However, the distribution of species with masses between 3 and 4 MDa is different, and unlike the situation with the lower dimer concentration, they do not appear to progress much with time. There is a peak at around 3.1 MDa accompanied by a broad distribution of species with masses between 3.1 and 4.0 MDa. The peak of this distribution is at around 3.7 MDa.

CDMS spectra showing the assembly of 20 μM dimer in 210 mM ammonium acetate and 5 μM dimer in 510 mM ammonium acetate can be found in Figures S1 and S2 of the Supporting Information. Increasing the dimer concentration from 20 μM under low salt conditions (210 mM ammonium acetate) produced assembly mass distributions similar to those shown in Figure 1B (i.e., highly abundant $T = 4$ capsid with a low abundant feature around 3.1 MDa). In addition, assembly with 5 μM dimer in 510 mM ammonium acetate retains trends similar to Figure 1C: intermediates around 3 MDa are greatly enhanced despite the low protein concentration.

Another notable feature of the results shown in Figure 1D is the substantial high mass tail that extends from the peak at around 4.1 MDa. A closer inspection of all the results in Figure 1 shows that there are high mass tails on the 4.1 MDa peak under all conditions, indicating capsid overgrowth. The tail in Figure 1D, where reactions have high Cp149 and salt concentrations, is the most prominent, extending to almost 5 MDa.

We previously used CDMS to detect kinetically trapped intermediates from the assembly of Cp149.[36] In that work, assembly was initiated by in 1 M sodium chloride, and after 24 h reactions were dialyzed for 24 h against 300 mM ammonium acetate for analysis. These conditions provided long times

DOI: 10.1021/jacs.8b01804
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

under aggressive assembly conditions for complexes to evolve and then long times under mild conditions for the less stable complexes to dissociate or rearrange. We observed a broad distribution of species with masses between 3.2 and 4.0 MDa with prominent features at 3.5 and 3.7 MDa. We do not observe the same prominent features here. However, we do observe a similar broad distribution of intermediates with masses between 3 and 4 MDa.

Figure 2 shows a higher resolution CDMS spectrum for the mildest assembly conditions in Figure 1 (panel A: 5 μM dimer

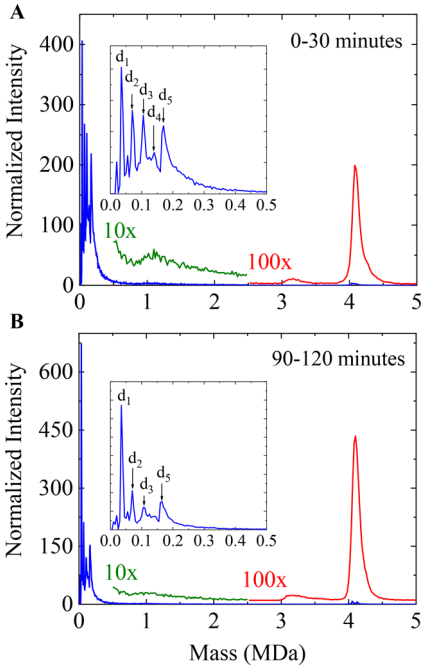


Figure 2. Higher resolution CDMS spectrum of a facile assembly reaction. (A) CDMS spectrum for the first 30 min and (B) for 90−120 min after initiation of the assembly reaction with a dimer concentration of 5 μM in 210 mM ammonium acetate. The blue spectrum was generated with 5 kDa bins. The insets in (A) and (B) show an expanded view of the low mass region up to 0.5 MDa where dimer (d_1) and oligomers up to five dimers (d_5) are resolved. The overlaid green and red spectra have been scaled up by factors of 10× and 100×, respectively, and are vertically offset. They were generated with 20 kDa bins.

in 210 mM ammonium acetate). Figure 2A compiles the first 30 min of the reaction. The spectrum was generated by binning all ions measured in the first 30 min into 5 kDa bins. The inset shows an expanded view of the low mass region where peaks due to the dimer and oligomers of up to five dimers are resolved. The oligomers are denoted by d_n, where n represents the number of dimers present. The green line shows ions with masses between 0.5 and 2.5 MDa binned into 20 kDa bins, enlarged by a factor of 10. A broad low-intensity peak is evident at around 1.1 MDa. The red line shows ions above 2.5 MDa binned into 20 kDa bins, enlarged by a factor of 100. A small broad peak at around 3.2 MDa is evident.

Figure 2B shows the corresponding spectrum for the 90−120 min time interval. Peaks are resolved for dimer to $(dimer)_5$ (see inset in Figure 2B). At this later time interval, the peaks corresponding to $(dimer)_2$ through $(dimer)_5$ have decreased significantly relative to the free dimer peak. The broad low-intensity feature observed at around ∼1.1 MDa in the 0−30 min spectrum has almost disappeared. The fact that this feature is broad, diminishes as the reaction proceeds, and the capsid peak grows, suggests that the peak is an ensemble of intermediates. Similarly, the small peak at around 3.1 MDa in the 0−30 min spectrum (Figure 2A), has broadened slightly, and shifted to slightly higher mass in the 90−120 min spectrum (Figure 2B), and over time it continues to shift to higher mass. The prominent peak at around 4.1 MDa has a high mass tail which continues to extend toward higher mass over the course of 0−30 min.

The spectra in Figure 2 are dominated by the peaks due to the dimer and dimer oligomers. The relative abundance of the capsid peak is very low. The low relative abundance results because the initial dimer concentration (5 μM) is only slightly larger than the pseudo-critical value (3.55 μM). After 24 h, the ratio of the number of dimers in the low mass species (<2.5 MDa) in the CDMS spectrum to the number of dimers in the high mass species (>2.5 MDa) is in reasonable agreement with ratio of dimer to capsid determined at the same time point with size exclusion chromatography.[43]

Assembly in 510 mM ammonium acetate generates a much higher abundance of species with masses between 3 and 4 MDa. Figure 3 shows a more detailed view of the species generated with an initial dimer concentration of 10 μM. The mass spectrum represented by the blue trace in Figure 3A was

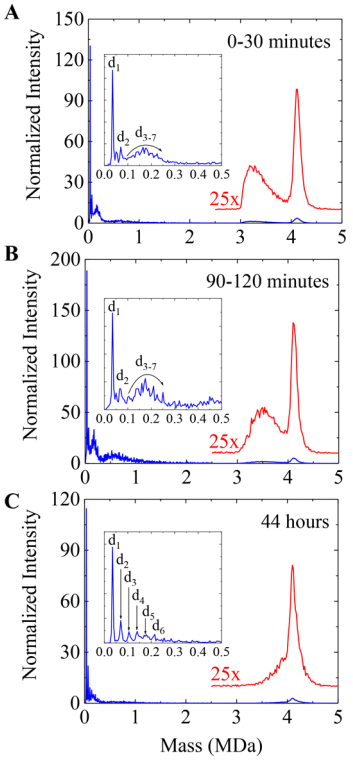


Figure 3. Higher resolution CDMS spectrum of an assembly reaction at strong association energy that generates many incomplete particles. (A) CDMS spectrum measured for the first 30 min, (B) 90−120 min, and (C) 44 h after initiation of an assembly reaction with a dimer concentration of 10 μM in 510 mM ammonium acetate. The blue spectra were generated with 5 kDa bins. The insets in (A−C) show an expanded view of the low mass region up to 0.5 MDa where dimer (d_1) and dimer oligomers are resolved. The overlaid red spectra were generated with 20 kDa bins and have been scaled up by a factor of 25×.

DOI: 10.1021/jacs.8b01804
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

generated by combining all ions from the first 30 min of the assembly reaction shown in Figure 1C into 5 kDa bins. The inset shows an expanded view of the low mass region up to 0.5 MDa where peaks due to the dimer and dimer oligomers are resolved. The red trace shows ions with masses between 2.5 and 5 MDa binned into 20 kDa bins, magnified by a factor of 25, and vertically offset. The inset in Figure 3A shows the most abundant peak in the spectrum which corresponds to dimer. There is a small amount of $(dimer)_2$, and a broad distribution of masses out to 0.25 MDa corresponds to dimer oligomers from $(dimer)_3$ to $(dimer)_7$; however, the larger oligomers are poorly resolved, partly due to their low abundances. Few species are observed between 1 and 3 MDa. A broad distribution exists between 3 and 4 MDa, with the apex of the distribution centered at around 3.2 MDa. The prominent peak at around 4.1 MDa is close to the mass expected for the $T = 4$ capsid. Figure 3B shows the spectrum for a time interval of 90−120 min. Above 3 MDa, the broad distribution has shifted to a higher mass, with the apex at around 3.5 MDa. After 44 h (Figure 3C), the broad peak between 3 and 4 MDa appears to have merged with the capsid peak at around 4.1 MDa. The capsid peak still has a low-mass tail, suggesting that some capsids are still struggling to complete. The inset in Figure 3C shows that the dimer and dimer oligomers up to $(dimer)_6$ are resolved at this time.

Figure 4 shows CDMS spectra measured over 5 days for the assembly of 20 μM dimer in 510 mM ammonium acetate; the conditions of this reaction are identical to those used for Figure 1D. The x-axis shows the mass binned into 20 kDa bins, the y-axis is the intensity normalized by peak area, and the z-axis represents the reaction time binned into 1 day intervals. After 1 day there is still a relatively sharp peak at 3.1 MDa (around 90 dimers) along with a broad, lower intensity distribution between 3.2 and 4 MDa. Over the next few days the sharp peak at 3.1 MDa decreases in abundance and appears to shift in mass toward the $T = 4$ capsid. This trend is consistent with the behavior observed for assembly under lower dimer and salt conditions. After 5 days, there is no longer a discernible peak at 3.1 MDa, and the intensity between 3 and 4 MDa has become much lower. These results indicate that the peak at 3.1 MDa is not due to the $T = 3$ capsid as the $T = 3$ capsid is stable in ammonium acetate. $T = 3$ capsids assembled in sodium chloride and dialyzed into ammonium acetate persist in solution for months.

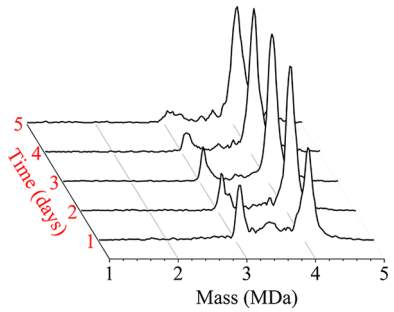


Figure 4. Time-resolved CDMS spectra showing the progression of capsid assembly over 5 days for a reaction mixture containing a total protein (dimer) concentration of 20 μM in 510 mM ammonium acetate. The earliest time interval (1 day) is at the front. The spectra were generated using 20 kDa bins.

DISCUSSION

With the Lower Salt Concentration There Are Efficient Downhill Pathways from Dimer to Capsid, and Few Intermediates Are Observed. The main features in the CDMS spectrum measured for assembly with the lower salt concentration (210 mM) are the dimer, small oligomers, and a relatively narrow peak at around 4.1 MDa attributed to the overgrown capsid.[43] For the lower dimer concentration (5 μM) there is also a broad peak centered on around 1.1 MDa (around 32 dimers) (see Figure 2A). This feature is most abundant in the spectrum for the earliest time point (see Figure 1A). Though its identity cannot be determined by CDMS, a compact circular patch of dimers is predicted to be most stable based on maximizing interdimer contacts and minimizing line tension; it is also predicted to be kinetically accessible.[15,49] The small oligomers and the feature at around 1.1 MDa behave like intermediates. They are observed early in the reaction and depleted as the reaction proceeds. The low abundance of species with masses between 2 and 4 MDa indicates that there must be a downhill pathway ensemble from the dimer and small oligomers to the overgrown capsid with no significant pause points along the way (i.e., a pathway with a steep potential energy gradient associated with efficient self-assembly[19,50]). The steep pathway probably involves intermediates that are fragments of the complete capsid. The small oligomers are prominent during assembly in agreement with predictions from master equations.[14,51]

With the Higher Salt Concentration around Half of the Assembly Products Are Trapped as Low Free Energy Intermediates with Masses between 3 and 4 MDa. When the salt concentration is increased to 510 mM, the first phase of the reaction is complete in less than a minute. However, only around half of the products of the initial assembly reaction are in the peak at around 4.1 MDa due to the overgrown capsid. This indicates that there is a facile pathway available. Most of the remaining balance is intermediates with masses between 3 and 4 MDa. With a dimer concentration of 10 μM, the 3−4 MDa intermediates gradually shift to higher mass during the 90 min time scale shown in Figure 1C. Inspection of Figure 3 shows that a significant amount of the dimer remains after the initial assembly reaction is complete. The pseudo-critical concentration for assembly of capsid under high salt conditions is around 0.7 μM.[43] Thus, it is unlikely that the stalled intermediates are due to low dimer concentration. Thus, species with masses between 3 and 4 MDa are not trapped by the absence of dimer needed to proceed to capsid. Instead, these species resist proceeding to the complete capsid because they are local free energy minima. It is anticipated that some of these species are off pathway due to puckering (suggested below), association of mismatched fragments, or trapped defects that increase the lifetime of the intermediates.

With Aggressive Assembly Conditions the 3−4 MDa Intermediates Are Kinetically Trapped. With a dimer concentration of 20 μM and high ionic strength, the 3−4 MDa intermediates do not progress during 90 min shown in Figure 1D. However, the dimer concentration after the initial assembly reaction is much lower than with the 10 μM initial dimer concentration and lower ionic strength and difficult to measure by CDMS. In high ionic strength, a large number of nuclei form early in the assembly reaction, and thus insufficient dimer is present to complete the capsid. In these circumstances, dissociation becomes a limiting step in the assembly reaction;

DOI: 10.1021/jacs.8b01804
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

some capsids must fall apart to generate dimers to be consumed by others. As shown in Figure 4, intermediates progress over a much longer time scale.

Stable Intermediates That Emerge at around 90 Dimers Are Not $T = 3$ Capsids. A striking feature of the results presented here is the abrupt emergence of intermediates at around 90 dimers (see Figure 1), close to the mass of a $T = 3$ capsid. However, the trapped intermediates slowly progress toward the $T = 4$ capsid, so their emergence at 90 dimers is not related to the $T = 3$ capsid. Assuming that facile pathways for capsid assembly involve intermediates that are incomplete capsid, there are several explanations for the accumulation of the intermediates at ≥90 dimers: (i) There is a subset of the capsid fragments that are more stable and progress more slowly under high salt conditions. However, these species should be observable at low protein concentrations but are not. (ii) There is another geometry that can only be accessed with 90 dimers involved. (iii) Instead of assembly one subunit at a time, higher concentrations of intermediates support association of several fragments resulting in nonicosahedral geometry with difficult-to-fill disclinations. However, there is no evidence of accumulation of the moderate sized intermediates needed to make this sort of chimera. (iv) Assembly is limited by defects in the lattice, making dissociation a limiting step in the formation of capsid.

Stable Intermediates That Emerge at around 90 Dimers Are Probably Not Incomplete Capsids. A plausible low-energy geometry for an incomplete $T = 4$ capsid missing up to around 30 dimers is a $T = 4$ capsid with a single, roughly circular hole. There is an energetic cost associated with the hole (the line tension) which results from missing dimer−dimer contacts. The line tension acts to minimize the length of the perimeter of the hole, making a single circular hole favorable to multiple holes and noncircular holes. A high symmetry intermediate with a single hole can be formed by removing a circular 30 dimer fragment from a $T = 4$ capsid. Because of its symmetry, addition of another dimer to this species would be less favorable because it would be bound by only two dimer−dimer contacts. This could provide an explanation for why the 90-dimer species is a prominent intermediate. However, it does not provide an explanation for why progression of all species between 3 and 4 MDa is slow, it does not provide an explanation for why progression is fast below 3 MDa where similar species could be devised, and it does not provide an explanation for why the intermediates become stalled when the ionic strength is increased. For the reverse reaction, it has been observed that a capsid with a singular hole is unlikely.[23]

We Speculate That Hole Collapse Accounts for the Emergence of Stable Intermediates at around 90 Dimers. It is known that a hole in a shell can collapse before the optimum number of subunits is present.[52] In this case, the energetic cost of distorting the shell to fill the hole is compensated by the loss of line tension. A plausible explanation for the sudden emergence of the trapped intermediates at 90 dimers is that this is the minimum size where it is energetically favorable for the hole to collapse. For a $T = 4$ capsid with a hole of around 30 dimers, distortion from a spherical geometry could allow the hole to heal. This geometry would be lower in energy than the undistorted $T = 4$ fragment but higher in energy than the complete $T = 4$, so it is still energetically favorable to complete the capsid. However, in order for the capsid to grow, it is now necessary to open up at least part of the sealed edge of the hole. There is an energetic cost associated with opening up the hole, leading to an activation energy for capsid growth. This would explain why the intermediates take so long to complete. By this logic, the reason the intermediates are only observed under high salt conditions is that the high ionic strength strengthens the dimer−dimer interactions so that the energy gained by sealing the hole is larger than the energetic penalty of distorting the capsid.

CONCLUSIONS

The results presented here indicate that assembly of the HBV capsid can follow multiple pathways with disparate time scales. The data support pathways that are not susceptible to trapping and pathways that are. Even successful assembly has evidence of multiples classes of assembly kinetics. Here, we observe a facile pathway from dimer to the overgrown capsid. This pathway occurs on an energy landscape that must be efficiently downhill. Under the lower salt conditions studied here almost all of the assembly reaction follows this route. The most prominent intermediates are small oligomers and a broad, low-abundance distribution with around 32 dimers. Few intermediates are observed between 2 and 4 MDa.

Under the higher salt conditions studied here the initial phase of the assembly reaction is fast, occurring in less than a minute. However, around half of the assembly is stalled at ≥90 dimers and then progresses much more slowly toward the capsid. This is potentially due to defects in the growing capsids that arise from stronger association energy. These defective particles can anneal to a low-energy state, but at a very slow pace. If this is the case, dissociation steps must be part of the process of a nascent capsid reaching the final product. Another plausible explanation for the stalled intermediates is that they result from hole closure where the capsid has distorted to close the hole due to the missing dimers. Increasing the salt concentration strengthens dimer−dimer interactions which favors hole closure. The abrupt appearance of the intermediates at around 90 dimers can be attributed to the presence of a minimum size for hole closure.

The data presented here are specific to HBV, but the picture of capsid assembly that arises from these data is probably general. The ensemble of assembly intermediates is exquisitely sensitive to concentration and association energy. It is striking that conditions with weak interdimer association, closer to physiological, lead to a funneled potential energy landscape with low downhill barriers whereas assembly conditions that favor strong association energy lead to a plethora of minima. Such landscapes have been visualized for a variety of self-organizing systems, including capsid assembly.[10,19,21] Theory has established that these landscapes are amenable to forming capsid shells of incorrect geometry and also that the probability of forming defective particles in a real system is high.[12] The persistence of these minima implies off path species—we favor a model for these intermediates where edges of hole in an incomplete lattice collapse and interact, blocking addition of the new subunits. An implication of this model is that incomplete capsids are extremely flexible.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.8b01804.

Figures S1 and S2 (PDF)

DOI: 10.1021/jacs.8b01804
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

AUTHOR INFORMATION

Corresponding Authors

*(A.Z.) E-mail azlotnic@indiana.edu.
*(M.F.J.) E-mail mfj@indiana.edu.

ORCID

Adam Zlotnick: 0000-0001-9945-6267
Martin F. Jarrold: 0000-0001-7084-176X

Notes

The authors declare the following competing financial interest(s): The authors, except A.Z., declare no competing financial interests. A.Z. is associated with a company that is developing antiviral compounds.

ACKNOWLEDGMENTS

We gratefully acknowledge the support of the NIH through Award Number 1R01AI118933.

REFERENCES

(1) Yeates, T. O.; Padilla, J. E. *Curr. Opin. Struct. Biol.* **2002**, *12*, 464−470.
(2) Padilla, J. E.; Colovos, C.; Yeates, T. O. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 2217−2221.
(3) Lai, Y. T.; Cascio, D.; Yeates, T. O. *Science* **2012**, *336*, 1129.
(4) King, N. P.; Sheffler, W.; Sawaya, M. R.; Vollmar, B. S.; Sumida, J. P.; André, I.; Gonen, T.; Yeates, T. O.; Baker, D. *Science* **2012**, *336*, 1171−1174.
(5) Lai, Y. T.; King, N. P.; Yeates, T. O. *Trends Cell Biol.* **2012**, *22*, 653−661.
(6) King, N. P.; Bale, J. B.; Sheffler, W.; McNamara, D. E.; Gonen, S.; Gonen, T.; Yeates, T. O.; Baker, D. *Nature* **2014**, *510*, 103−108.
(7) Lai, Y.-T.; Reading, E.; Hura, G. L.; Tsai, K.-L.; Laganowsky, A.; Asturias, F. J.; Tainer, J. A.; Robinson, C. V.; Yeates, T. O. *Nat. Chem.* **2014**, *6*, 1065−1071.
(8) Hsia, Y.; Bale, J. B.; Gonen, S.; Shi, D.; Sheffler, W.; Fong, K. K.; Nattermann, U.; Xu, C.; Huang, P. S.; Ravichandran, R.; Yi, S.; Davis, T. N.; Gonen, T.; King, N. P.; Baker, D. *Nature* **2016**, *535*, 136−139.
(9) Wales, D. J.; Miller, M. A.; Walsh, T. R. *Nature* **1998**, *394*, 758−760.
(10) Wales, D. J.; Bogdan, T. V. *J. Phys. Chem. B* **2006**, *110*, 20765−20776.
(11) Zandi, R.; Reguera, D.; Bruinsma, R. F.; Gelbart, W. M.; Rudnick, J. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 15556−15560.
(12) Fejer, S. N.; Chakrabarti, D.; Wales, D. J. *ACS Nano* **2010**, *4*, 219−228.
(13) Caspar, D. L. D.; Klug, A. *Cold Spring Harbor Symp. Quant. Biol.* **1962**, *27*, 1−24.
(14) Endres, D.; Zlotnick, A. *Biophys. J.* **2002**, *83*, 1217−1230.
(15) Zandi, R.; van der Schoot, P.; Reguera, D.; Kegel, W.; Reiss, H. *Biophys. J.* **2006**, *90*, 1939−1948.
(16) Nguyen, H. D.; Reddy, V. S.; Brooks, C. L. *Nano Lett.* **2007**, *7*, 338−344.
(17) Hagan, M. F.; Elrad, O. M. *Biophys. J.* **2010**, *98*, 1065−1074.
(18) Hagan, M. F. *Adv. Chem. Phys.* **2014**, *155*, 1−67.
(19) Wales, D. J. *J. Theor. Medicine* **2005**, *6*, 107−110.
(20) Wales, D. J. *Philos. Trans. R. Soc., A* **2005**, *363*, 357−377.
(21) Fejer, S. N.; James, T. R.; Hernández-Rojas, J.; Wales, D. J. *Phys. Chem. Chem. Phys.* **2009**, *11*, 2098−2104.
(22) Zlotnick, A. *J. Mol. Biol.* **1994**, *241*, 59−67.
(23) Lee, L. S.; Brunk, N.; Haywood, D. G.; Keifer, D. Z.; Pierson, E. E.; Kondylis, P.; Wang, J. C.-Y.; Jacobson, S. C.; Jarrold, M. F.; Zlotnick, A. *Protein Sci.* **2017**, *26*, 2170−2180.
(24) Fuerstenau, S. D.; Benner, W. H. *Rapid Commun. Mass Spectrom.* **1995**, *9*, 1528−1538.
(25) Keifer, D. Z.; Pierson, E. E.; Jarrold, M. F. *Analyst* **2017**, *142*, 1654−1671.
(26) Beck, J.; Nassal, M. *World J. Gastroenterol.* **2007**, *13*, 48−64.
(27) Nassal, M. *Virus Res.* **2008**, *134*, 235−249.
(28) Ning, X.; Nguyen, D.; Mentzer, L.; Adams, C.; Lee, H.; Ashley, R.; Hafenstein, S.; Hu, J. *PLoS Pathog.* **2011**, *7*, e1002255.
(29) Ceres, P.; Zlotnick, A. *Biochemistry* **2002**, *41*, 11525−11531.
(30) Kenney, J. M.; von Bonsdorff, C. H.; Nassal, M.; Fuller, S. D. *Structure* **1995**, *3*, 1009−1019.
(31) Dryden, K. A.; Wieland, S. F.; Whitten-Bauer, C.; Gerin, J. L.; Chisari, F. V.; Yeager, M. *Mol. Cell* **2006**, *22*, 843−850.
(32) Zlotnick, A.; Johnson, J. M.; Wingfield, P. W.; Stahl, S. J.; Endres, D. *Biochemistry* **1999**, *38*, 14644−14652.
(33) Uetrecht, C.; Versluis, C.; Watts, N. R.; Roos, W. H.; Wuite, G. J. L.; Wingfield, P. T.; Steven, A. C.; Heck, A. J. R. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 9216−9220.
(34) Uetrecht, C.; Barbu, I. M.; Shoemaker, G. K.; van Duijn, E.; Heck, A. J. R. *Nat. Chem.* **2011**, *3*, 126−132.
(35) Holmes, K.; Shepherd, D. A.; Ashcroft, A. E.; Whelan, M.; Rowlands, D. J.; Stonehouse, N. J. *J. Biol. Chem.* **2015**, *290*, 16238−16245.
(36) Pierson, E. E.; Keifer, D. Z.; Selzer, L.; Lee, L. S.; Contino, N. C.; Wang, J. C. Y.; Zlotnick, A.; Jarrold, M. F. *J. Am. Chem. Soc.* **2014**, *136*, 3536−3541.
(37) Zlotnick, A.; Ceres, P.; Singh, S.; Johnson, J. M. *J. Virol.* **2002**, *76*, 4848−4854.
(38) Contino, N. C.; Jarrold, M. F. *Int. J. Mass Spectrom.* **2013**, *345*−*347*, 153−159.
(39) Contino, N. C.; Pierson, E. E.; Keifer, D. Z.; Jarrold, M. F. *J. Am. Soc. Mass Spectrom.* **2013**, *24*, 101−108.
(40) Pierson, E. E.; Keifer, D. Z.; Contino, N. C.; Jarrold, M. F. *Int. J. Mass Spectrom.* **2013**, *337*, 50−56.
(41) Pierson, E. E.; Contino, N. C.; Keifer, D. Z.; Jarrold, M. F. *J. Am. Soc. Mass Spectrom.* **2015**, *26*, 1213−1220.
(42) Keifer, D. Z.; Shinholt, D. L.; Jarrold, M. F. *Anal. Chem.* **2015**, *87*, 10330−10337.
(43) Lutomski, C. A.; Lyktey, N. A.; Zhao, Z.; Pierson, E. E.; Zlotnick, A.; Jarrold, M. F. *J. Am. Chem. Soc.* **2017**, *139*, 16932−16938.
(44) Crowther, R. A.; Kiselev, N. A.; Bottcher, B.; Berriman, J. A.; Borisova, G. P.; Ose, V.; Pumpens, P. *Cell* **1994**, *77*, 943−950.
(45) Wingfield, P. T.; Stahl, S. J.; Williams, R. W.; Steven, A. C. *Biochemistry* **1995**, *34*, 4919−4932.
(46) Stannard, L. M.; Hodgkiss, M. J. *J. Gen. Virol.* **1979**, *45*, 509−514.
(47) Selzer, L.; Katen, S. P.; Zlotnick, A. *Biochemistry* **2014**, *53*, 5496−5504.
(48) Zlotnick, A.; Cheng, N.; Conway, J. F.; Booy, F. P.; Steven, A. C.; Stahl, S. J.; Wingfield, P. T. *Biochemistry* **1996**, *35*, 7412−7421.
(49) Moisant, P.; Neeman, H.; Zlotnick. *Biophys. J.* **2010**, *99*, 1350−1357.
(50) Johnston, I. G.; Louis, A. A.; Doye, J. P. K. *J. Phys.: Condens. Matter* **2010**, *22*, 104101.
(51) Keef, T.; Micheletti, C.; Twarock, R. *J. Theor. Biol.* **2006**, *242*, 713−721.
(52) Luque, A.; Reguera, D.; Morozov, A.; Rudnick, J.; Bruinsma, R. *J. Chem. Phys.* **2012**, *136*, 184507.

DOI: 10.1021/jacs.8b01804
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX