HHS Public Access

Author manuscript

Phys Biol. Author manuscript; available in PMC 2019 June 06.

Published in final edited form as:

Phys Biol. ; 15(5): 056005. doi:10.1088/1478-3975/aac194.

Molecular Jenga: The Percolation Phase Transition (Collapse) in Virus Capsids

Nicholas E. Brunk[1,2], **Lye Siang Lee**[1], **James A. Glazier**[2,3], **William Butske**[4], and **Adam Zlotnick**[1]

[1]Department of Molecular and Cellular Biochemistry, Indiana University, Bloomington, IN 47405, USA

[2]Department of Intelligent Systems Engineering, Indiana University, Bloomington, IN 47405, USA

[3]The Biocomplexity Institute, Indiana University, Bloomington, IN 47405, USA

[4]Department of Mathematics, Rose-Hulman Institute of Technology, Terre Haute, IN 47803, USA

Abstract

Virus capsids are polymeric protein shells that protect the viral cargo. About half of known virus families have icosahedral capsids that self-assemble from tens to thousands of subunits. Capsid disassembly is critical to the lifecycles of many viruses yet is poorly understood. Here, we apply a graph- and percolation theory to examine the effect of removing capsid subunits on capsid stability and fragmentation. Based on the structure of the icosahedral capsid of Hepatitis B Virus (HBV), we constructed a graph of rhombic subunits arranged with icosahedral symmetry. Though our approach neglects dependence on energetics, time, and molecular detail, it quantitatively predicts a percolation phase transition consistent with recent *in vitro* studies of HBV capsid dissociation. While the stability of the capsid graph followed a gradual quadratic decay, the rhombic tiling abruptly fragmented when we removed more than 25% of the subunits, near the percolation threshold observed experimentally. This threshold may also affect results of capsid assembly, which also experimentally produces a preponderance of 90mer intermediates, as the intermediate steps in these reactions are reversible and can thus resemble dissociation. Application of percolation theory to understanding capsid association and dissociation may prove a general approach to relating virus biology to the underlying biophysics of the virus particle.

1. Introduction

In the game of Jenga, one must remove pieces from a tower of wooden blocks without causing the tower to collapse. In percolation theory, the definition of the percolation threshold is the point where removed elements break a connected lattice into disconnected fragments, analogous to the collapse of the blocks in Jenga, or in a wooden block mechanical demonstration of the percolation threshold [1]. Here we are interested in modeling removal of subunits from a viral capsid, the protein shell of a virus, anticipating its collapse. In about half of known virus families, the capsid is arranged with icosahedral symmetry and thus has 60 equivalent asymmetric units, where an asymmetric unit may be comprised of one or more proteins. The role of the capsid in the virus life cycle is to protect, deliver, and release the viral cargo. To perform this function *in vivo*, capsids must readily

assemble and, for many viruses, readily disassemble. Computational models have been used to understand the process of assembly. Models can generally be grouped as master equations and coarse-grained molecular dynamics [2]. Robust assembly can be accomplished by addition of one subunit at a time even when intersubunit association energy is weak because subunits are multivalent. To minimize kinetic trapping, assembly may depend on a nucleation step or be subject to allosteric regulation [3–10]. Irreversible assembly or strong association energy can lead to trapped errors [11–13].

For example, the icosahedral capsid of Hepatitis B Virus (HBV) has become an important model system for studying capsid biophysics. HBV is also an important pathogen as approximately 240 million people have chronic HBV and it contributes to 780,000 deaths each year. In vivo, HBV assembles in the cytoplasm and disassembles prior to or during nuclear entry [14, 15]. The HBV capsid is composed of 120 copies of the homodimeric core protein (Figure 1); a small fraction of capsids are built of 90 dimers. In the 120-dimer capsid, subunits are arranged with T=4 icosahedral symmetry (Figure 1A), while the 90-dimer capsid has T=3 symmetry. Following the Caspar-Klug theory of quasiequivalence, a T=4 icosahedron has twelve pentameric vertices and thirty hexameric vertices [16]. The dimeric subunits make four intersubunit connections and either extend from a pentamer to a hexamer or between two hexamers. Experimental estimates place dimer-dimer interaction energy at about 6 k_BT (−3.6 kcal/mol), which favors assembly under near physiological conditions only because dimers are tetravalent. Thus, in vivo and in vitro, conditions can be found where assembly is the lowest energy state. In spite of the relative fragility of association, even when the capsid form is not the thermodynamically favored state, HBV capsids are remarkably stable, persisting for months without substantial exchange of subunits and surviving solution conditions where they are expected to dissociate [17, 18]. As one might expect, assembly models can be run in reverse to model capsid disassembly, with successful replication of hysteresis observed *in vitro* [17,18]. Recently, the HBV core protein has become a target for development of antiviral therapeutics; of note, these antiviral molecules favor capsid assembly [19–22].

Our understanding of disassembly remains limited. Our interest in the mechanism of dissociation was piqued by recent work where we studied dissociation of subunits from in vitro assembled HBV capsids [23]. In these solution experiments, subunits were removed from intact capsids by gentle treatment with a denaturant, leaving behind incomplete capsids. In some experiments, a specified fraction of the subunits within a capsid were chemically crosslinked so that only uncrosslinked subunits could be removed. With both uncrosslinked capsids and partially crosslinked capsids, using single particle mass spectrometry and nanofluidic approaches, we observed complexes between intact 120-mers and partially complete 90-dimer fragments. We observed essentially no species with fewer than 90 subunits [23]. It is possible that any smaller intermediates are insufficiently stable and/or too rare to be detected. Cryo-electron microscopy showed that incomplete particles, crosslinked or uncrosslinked, had no obvious features, e.g. a single hole, that distinguished them from T=4 particles. However, the incomplete particles were “live polymers” as free dimer, fluorescently labeled, and could be added to partially complete capsids to rebuild T=4 particles. The array of incomplete capsids seen during disassembly closely matches those seen in assembly reactions [24–26], at least in terms of the masses of the observed particles.

This result leads us to question what features there might be in a T=4 120-mer capsid that leads to accumulation of a species missing 30 subunits.

To mimic disassembly reactions and investigate the underlying physics, we have developed a simplified graph representation of the capsid and randomly removed subunits from it in Monte Carlo simulations (Figure 1B, C). The graph description of capsid disassembly has important similarities to examination of percolation in regular lattices. For a very large lattice, subtraction of random subunits results in an abrupt phase transition or percolation threshold, between a single connected graph and many small disconnected subgraphs [27–32]. Three important peculiarities distinguish a capsid from most general models of regular graphs. First, a capsid is a thin-walled sphere whose surface can be considered two-dimensional. Second, capsids are assembled from a relatively small number of subunits, usually tens to hundreds, leading to a more gradual phase transition than is seen for large graphs. Third, comparison of models and data may not be straightforward as the simplified graph vertices represent actual proteins, which may change conformation during assembly, have complex interactions with one or more subunits, and may have different interactions at different interfaces.

We first applied percolation theory to explain experimental observations of dissociation [23]; here we examine the implications of percolation theory in greater depth. The character of a virus-like graph – the small number of subunits arranged in a two dimensional array – impact its fragmentation. To examine disassembly, we have adopted a coarse-grained simplification of subunits. We observed that capsid-like graphs break into fragments over a relatively narrow range of removed subunits, but not at the precise transition of an infinitely large graph given the small number of subunits composing the virus. The predicted percolation threshold of about 25% for HBV tiling (i.e. fragmentation if missing more than 30 of 120 “dimers” in a capsid) agreed well with experimental observations. We found that as we randomly remove subunits, the stability of the capsid, proportional to the number of intersubunit contacts, falls quadratically. The two-dimensional nature of the graph, that is the lack of three-dimensional connections between subunits, makes it particularly susceptible to fragmentation. From a biological perspective, these results imply that dissociation will require removal of several subunits for a capsid to release its cargo – a single deleted subunit may be insufficient to induce particle collapse. Interaction with the viral cargo, such as nucleic acid, will further impede dissociation. The ability to predict capsid fragmentation while neglecting energetics and time-dependence demonstrates utility of a simplified percolation theoretical modeling technique. Discrepancies between the model and *in vitro* observations may provide insight into the underlying capsid biophysics.

2. Definitions and Methods

A graph G is comprised of vertex and edge sets $V(G)$ and $E(G)$. The order of a graph G is the number of vertices V_0. For comparison of the capsid graph to viruses, each vertex is equivalent to a subunit. The number of edges between vertices in G is $E(0)$. A graph G is k-regular if the number of adjacent vertices to each vertex v_j is k for all vertices in the graph. The graph of the HBV capsid in Figure 1 has $V_0 = 120$ and is 4-regular. In Figure 1, the

graph and capsid are dual to one another; by definition, vertices represent subunits and edges represent interstitial bonds or contacts.

We define $G(i)$ to be the subgraph that results from deleting i vertices and their associated edges. The fraction of deleted vertices is thus $f_V^d(i) \equiv \left(\frac{i}{V_0}\right)$ and the fraction of vertices remaining in $G(i)$ is $\rho_V(i) \equiv (i - f_V^d(i))$. Vertex fraction remaining $\rho_V(i)$ corresponds to the "lattice site" occupation probability $p_V(i)$ in percolation theory [30, 32]. We implemented disassembly and analysis in *Mathematica 10.3.0* (Wolfram Research).

In section 3.1 we concern ourselves with quantities associated with $G(i)$. The number of remaining edges constituting $G(i)$, denoted $E(i)$, corresponds to the number of intersubunit contacts in the system (equation 2.1). Note that bars indicate the number of elements within the set.

$$E(i) \equiv |E(G(i))| \quad 2.1)$$

In section 3.2 we investigate a more consequential quantity of interest for partially disassembled viruses: $b_0(i)$ is the mean number of connected components, i.e. the number of polymeric fragments or clusters. The number of fragments $b_0(i)$ from a given Monte Carlo simulation of disassembly is the 0^{th} Betti number in topology: the number of connected components for each specific instance of $G(i)$ [33]. Because $G(i)$ may be any one of an ensemble of graphs, we estimate $\bar{b}_0(i)$ as the arithmetic mean from many simulations for a given value of i. If $b_0(i) = 1$, a partially disassembled graph $G(i)$ is connected. In section 3.3 we use $\bar{b}_0(i)$ to determine the percolation threshold numerically via Monte Carlo. When $\bar{b}_0(i) \geq 2$, fragmentation has occurred.

In section 3.4 we use Equation 2.2 to characterize partially disassembled graphs $G(i)$ by their local clustering coefficients, $C(v_j)$ [34]. $C(v_j)$ is the per vertex v_j, that is the actual number of edges $\in_j$ between vertex v_j and its $k(v_j)$ neighbors divided by the maximum possible number of edges for all adjacent vertices:

$$C\left(v_j\right) = \frac{\left|\left(v_j, v_k\right) \in E(G) \middle| v_j, v_k \in N\left(v_i\right)\right|}{\left|\left\{\left(v_j, v_k\right) \middle| v_j, v_k \in N\left(v_i\right)\right\}\right|} = \frac{\varepsilon_j}{\left(\frac{k\left(v_j\right) * \left(k\left(v_j\right) - 1\right)}{2}\right)} \quad 2.2)$$

For an isolated vertex and pairs of vertices lacking other neighbors, the numerator, denominator, or both equal zero. We follow the convention of setting $C(v_j)$ for these vertices to zero [34]. In partially isolated vertices $C(v_j)$ is 1. For vertices in intact particles, or where the local environment resembles an intact particle, the $C(v_j)$ is 1/3 (Table 1).

3. Results

3.1 Graph edge count (capsid stability) decays quadratically with the number of vertices removed

To mimic the experiment with partially crosslinked capsids [23], we model removal of i subunits from a capsid by removing random vertices from a graph. While this approach to dissociation does not preferentially remove subunits based on Boltzmann weighting of stability, it does not require that we make assumptions that any particular set of subunits are more labile. As the number of removed vertices increases, the edge count, or number of contacts between subunits, $E(i)$, decreases gradually (Figure 2). The number of intersubunit contacts in a capsid is essentially a measure of the association energy of the capsid, thus the edge count is an index of capsid stability. Empirically, the edge count of partially disassembled graph $G(i)$ decays quadratically with i, the number of vertices removed. For all graphs, k-regular and random, that we have tested:

$$E(i) \approx \bar{E}(i) = \frac{\bar{k}_0}{2|V(G)|}i^2 - \bar{k}_0 i + |E(G)| = E(0) \quad - \bar{k}_0 i + \frac{\bar{k}_0}{2V_0}i^2 \qquad 3.1)$$

Recall, V_0 and $E(0)$ are the initial number of vertices and edges, respectively, and $\bar{k}_0$ is the initial number of edges per vertex (vertex degree or subunit valency). For a k_0-regular graph such as HBV (Figure 1) and small values of i, there is a low probability any two deleted vertices are adjacent, in which case the linear term dominates. The quadratic coefficient accounts for deleted vertices that are adjacent to one another, decrementing edges removed per unit vertex removed at greater values of i. Surprisingly, we find that Equation 3.1 applies to all graphs by substituting the mean vertex degree $\bar{k}_0$, it is not necessary for $G(0)$ to be k_0-regular. Using a number of test graphs, including metabolic networks we observe that not all subunits need to have the same valency in order for the equation to apply (Figure 2).

Equation 3.1 may be simplified in terms of the vertex set where $V(i) = |V(G)| - i = V_0 - i$:

$$\bar{E}(i) = \frac{\bar{k}_0}{2|V(G)|}(i - |V(G)|)^2 = \frac{\bar{k}_0}{2V_0}V(i)^2 \qquad 3.2)$$

Normalization of Equation 3.2 by $E_0 = \frac{\bar{k}_0 V_0}{2}$ and rearrangement suggests that the fraction of remaining edges $p_E(i)$ is scale invariant, in both $E(0)$ and V_0, equal to the fraction of remaining vertices $p_V(i)$ squared:

$$\frac{\bar{E}(i)}{E(0)} = \frac{V(i)^2}{V_0^2} \; or \; p_E(i) = p_V(i)^2 \qquad 3.3)$$

The decrease in edge count agrees well with the empirical equation, even for a single random subset of removed vertices per point (Figure 2). After a sufficient number of Monte Carlo replicates, the average converges to the predicted value (inset Figure 2A). Note that this predicts a fairly gradual decrease in the stability of the partially disassembled fragment(s) (*i*). This gradual decrease in ensemble stability is not expected given the abrupt fragmentation observed *in vitro* and indicates energetics are not the factor determining fragmentation.

We considered several alternative strategies for removing subunits. With an energy of $6k_BT$, typical for interaction energy between virus subunits, there is a strong bias for removing subunits from a single defect. The effect is to favor a single growing hole in the capsid, where loss of each subunit leads to on average a loss of two intersubunit contacts (Figure 2A). This so-called pacman approach seems like the intuitively the obvious way for a virus to lose subunits, however, experimentally this prediction is not borne out [23]. Microscopy shows that most do not have one contiguous hole during disassembly.

3.2 Fragmentation: statistically expected value of the cluster count, $\bar{b}_0(\mathbf{i})$

An assessment of fragmentation, the mean number of distinct fragments $\bar{b}_0(i)$, is an informative property of the partially disassembled graph $G(i)$ (Figure 3). This is known as the zeroth Betti number in topology. For a small number of vertices i deleted randomly from the graph, the graph remains intact. For HBV, as i approaches 30, the average number of fragments as a function of deleted vertices, $\bar{b}_0(i)$, rapidly increases, signifying a breakdown in the integrity of the capsid. Empirically, the fragmentation curve $\bar{b}_0(i)$, resembles the probability density function of the β-distribution, the so-called 'β-curve' $\mathscr{B}(\alpha,\beta)$. $\bar{b}_0(i)$ can be approximated as $\mathscr{B}(\alpha,\beta)$ with shape parameters α, β, and normalization constant c_n, which can be fit to $b_0(i)$ in terms of the fraction of vertices deleted $f_V^d(i) \equiv \left(\frac{i}{V_0}\right)$, and the fraction of vertices remaining $p_V(i) \equiv 1 - f_V^d(i)$:

$$\bar{b}_0(i) \approx \mathscr{B}\left(i, b_0(0), c_n, \alpha, \beta\right) \equiv b_0(0) - c_n f_V^d(i)^\alpha p_V(i)^\beta \quad 3.4)$$

where $b_0(0) = 1$ is the fraction of clusters in the initial graph.

Even for a 4-regular graph modeled on HBV, the breakdown of the graph in any given Monte Carlo simulation is highly variable. The standard deviations for $\bar{b}_0(i)$ show that – while abrupt –there is not a step-wise phase transition when subunits are removed randomly; this is due to the small size, it has relatively few subunits. Due to a significant standard deviation, observation of the β-curve requires multiple replicates.

3.3 Application of percolation theory to capsid disassembly

In a given partially disassembled graph $G(i)$, $b_0(i)$ is the number of fragments. We define $\bar{P}_{con}(i)$ as the probability that a partially disassembled graph will be connected. The

percolation threshold of the graph, p_{V_c}, is the fraction of sites that are filled p_V at the maximum of the first derivative of $\bar{P}_{con}(i)$ with respect to i (Figure 4) [30]. In practical terms, this definition of the percolation threshold occurs when about half of the graphs in a series of simulations have broken into two or more fragments $\bar{b}_0(i) \geq 2$.

For the 120 vertex HBV-based graph the percolation threshold of $p_V \approx 0.74$ occurs at $i_c \approx 31$ missing subunits. This value agrees well with solution experiments where we observed intermediates with up to, but no more than, 30 missing dimers. This threhsold is also observed in assembly experiments. An interesting divergence from experimental observation is that Monte Carlo simulations do show the presence of some intact graphs missing more than 30 subunits (Figure 4).

Graph topology determines resistance to fragmentation [32]. The HBV graph is planar, a 2-dimensional tiling over a spherical surface, in general a random graph is non-planar. As seen in Figure 4, because HBV is restricted to 2-dimensional contacts between neighboring subunits, it fragments more readily than a typical random graph. The percolation threshold for HBV is $p_{V_c} \approx 0.74$, higher than the average for multiple non-planar random, 4-regular graphs for which $p_{V_c} = \frac{120 - 36}{120} \approx .70$. This observation leads us to propose that planar graphs, and by analogy empty capsids, will fragment with fewer vertices removed. Similarly, figure 5 shows that for a given number of elem ents removed (i), HBV fragments more than the corresponding random non-planar graph ($\bar{b}_0^{HBV}(i) > \bar{b}_0^r(i)$). The susceptibility of planar graphs to fragmentation resembles the “extreme vulnerability of interdependent, spatially embedded” geographic networks [35].

Monte Carlo disassembly was also used to predict the size distribution of components when a graph fragmented. $\bar{P}_j(i)$ is defined as the probability of a connected graph of size j occuring in $G(i)$. For example, when a population of capsids are each missing $i = 10$ subunits, all of the capsids are perforated but intact and $P_{con} = \bar{P}_{110}(10) \approx 1$ (Figure 6a). When $i = 20$, 85% of the capsids remain intact. For $i = 30 \approx i_c$, close to the percolation threshold, 35% of capsids remain intact. For $i = 40$, only 5% of capsids remain intact. Over a population of capsids there will be a distribution of values of missing subunits so that when i is considerably above the percolation threshold we observe a binomial distribution of graph sizes(Figure 6b, see the magenta distribution for mean $i = 10$). When i is above the percolation threshold $i \geq i_c$, the number of intact graphs drops, but there may still be intact particles. We also observe in Figure 6 that the HBV-like graph, like most graphs, rarely fragments into two large clusters such that the probability intermediate values of j are very low, a hallmark of percolation theory consistent with solution studies.

3.4 Clustering coefficients, C(n)

To analyze the distribution of dissociation products, we consider the local stability of the HBV lattice as we remove subunits (Equation 2.3). In the partially disassembled HBV capsid, there exist three possible local subunit arrangements (Figure 1C). While useful as a

classification scheme indicating local stability, the value of the clustering coefficient itself is not proportional to stability. The local clustering coefficient for all vertices in the fully assembled capsid $G(0)$ is $C\left(v_j\right) = \frac{1}{3}$. Deviation from this indicates the local integrity of the cluster a vertex belongs to. Missing neighbors produce either a locally stable $C(v_j) = 1$ triangular arrangement (with each dimer bound to one another) or a $C(v_j) = 0$ arrangement with the subunits either in a weak, linear chain or isolated with no neighbors at all. Removal of a single subunit often fragments a linear chain (Figure 1C in red) but does not fragment a triangular arrangement (in green). The local clustering coefficient does not differentiate between local geometry as both fivefold and sixfold vertices are tetravalent (with the same topology). Figure 7 shows the fraction of subunits in capsid-like arrangements falls and the fraction of subunits in the easily fragmented linear configuration increases dramatically as vertices are removed. The $C(v_j) = 0$ pool of isolated subunits and linear chain arrangements becomes increasingly dominant as many vertices are removed, indicating the lack of any appreciable clusters. Near the percolation threshold, the number of vertices in the triangular configuration increases. By analogy, capsids held together by weak linear chains when $i \geq 30$ will be more prone to proceed with dissociation, which suggests why fragmentation is more abrupt in solution than in a static model.

4. Discussion

Dissociation of the capsid has biological, biochemical, and physical implications. In the lifecycle of many viruses, the dissociation and release of the viral genome is every bit as important as the assembly and packaging of the genome. So how does a virus capsid fall apart? Our initial hypothesis was that removal of a first subunit, or cluster of subunits, would lead to a catostrophic collapse [18]. The basis of this idea was in energetics and their effect on kinetics, that removing a subunit fully surrounded by neighbors would lead to a hole surrounded by many subunits held in by fewer contacts, leading the capsid to unravel. However, experimental studies showed that removing subunits from partially crosslinked and uncrosslinked HBV capsids resulted in accumulation of intermediates between 100% and about 75% of a complete capsid and very few smaller intermediates [23]. Microscopy suggests removal of subunits may not occur in one growing hole, but may be more random.

Remarkably, the same distribution of intermediates seen during idsassembly is also seen during assembly [24,25,36]. The study described in this paper suggests that there is a physical basis for the (meta)stability of a partially complete particle. We observe that removing more than the percolation threshold of $i_c \approx 31$ subunits will result in fragmentation of the intermediate into smaller, and presumably more labile, species. We suggest that the percolation threshold concept will be common to other spherical viruses, although the value of the threshold will change with the geometry of the lattice. We can not exclude the possibility of a biochemical basis peculiar to the T=4 HBV capsid for the stability of an incomplete particle, but it is not necessary to recapitulate the observed behavior.

We did not fully explore the mechanisms for dissociation. In this study, for simplicity, we considered the HBV-like graph to be built of 120 dimer-like rhombic subunits where all intersubunit interactions are equivalent. In HBV this seems a fair first approximation as

symmetrical intermediates such as trimers, pentamers, and hexamers do not accumulate during assembly [37,38]. Energetically, the difference between removing subunits randomly or preferentially from a growing hole described a relatively narrow envelope (Figure 2). Of course, in some viruses there are clearly stronger interactions between subunits observed experimentally (e.g. cowpea chlorotic mottle virus [39], Phi 29 [40], or minute virus of mice [41]) or predicted computationally [42,43]. A variegated pattern of stronger and weaker interactions is likely to generate a pattern of dissociation specific to a given virus that may affect stability and fragmentation. Antiviral compounds that modify the strength of intersubunit interactions [44] can also be predicted to modify dissociation.

Application of Monte Carlo simulations to dissociation of a capsid by removing one subunit at a time is a practical realization of percolation theory. Such simulations are used to determine percolation thresholds, the points at which the graphs fall apart, for the disassembly of arbitrary graphs. There are features of a virus capsid that impact percolation and stability:

(i) Capsids are small, often less than a few hundred subunits. Small graphs are expected to have relatively broad percolation transitions, as shown for both a capsid and random 120-subunit graphs in Figure 4. An examination of percolation in larger icosahedra saw a similar effect [45]. For very large graphs, one expects a more abrupt phase transition. Observation of “drilling percolation,” where the experimentalist literally drilled a series of holes through wooden blocks and characterized these observations with simulations that showed a similar effect as the relative size of the block increased [1].

(ii) Capsid lattices are two dimensional lattices stretched over a three dimensional solid, which limits the number of subunits and intersubunit interactions, making the graph more prone to fragmentation (Figure 5). Atomic force microscopy experiments investigating capsid indentation are consistent with this prediction. A prediction of this hypothesis, supported by experiment, is that a particle with nucleic acid crosslinking its inner surface would be expected to be far more resilient due to its non-planar topology [46–50].

(iii) Caspar-Klug quasi-equivalence dictates that interactions between subunits are not identical. In most cases, capsids have twelve fivefold vertices with a Caspar-Klug arrangement of surrounding sixfold vertices [51]. A more advanced model may consider different weighting factors for removal of subunits and stability of the remaining graph. Indeed, kinetics could also be factored into a more sophisticated model of dissociation.

(iv) Capsid subunits are made of protein(s), which have complex geometry. Interactions between subunits may be based on contact between two relatively featureless surfaces, require detailed complementarity, or require extended peptide chains to wrap around one another.

This study addresses the first two of these points.

The predicted percolation threshold for T=4 HBV capsids (120 dimers) was in agreement with solution experiments, except we do not experimentally observe the predicted continuum of large intermediates with fewer than 90 subunits [23,25,37]. For this purpose, we will define a large intermediate as having more than 60 subunits, which should be easily detected by single particle methods. Experimental observations of HBV dissociatiom intermediates were dominated by complexes of ≥90 dimers. Curiously, in assembly experiments we also observed a substantial pool of intermediates of ≥ 90 dimers [24,25]. These convergent experimental observations suggest that intermediates consisting of less than 90 dimers are particularly labile during disassembly or are consumed to make large species during assembly. We suggest the basis for the absence of 60- to 90-subunit intermediates arises from their lability, a kinetic effect. We note that throughout the dissociation simulation that the stability of the remaining graph continuously falls with i (Figure 2), which would be expected to progressively destabilize intermediates [8,52]. Furthermore, as i approaches and exceeds the percolation threshold, graphs of many intermediates are held together by a single intersubunit contact (Figure 7); such intermediates are unlikely to persist for long in solution.

Another surprise was that in experiments, we observed that the fraction of intact capsids was greater than our model predicted for a normal distribution of removed subunits [23]. We suspect this arose because we removed random subunits without regard to their neighbors; a pacman approach [53] of removing contiguous subunits and the requirement that two or more subnits be removed together partially addresses this discrepancy. The effects of different dissociation paths will be systematically investigated in the future. We will also need to take into account that intersubunit contacts are not necessarily equal. A T=4 HBV capsid has four quasi-equivalent environments with non-equivalent interactions; the fivefold interactions seem notably weaker than others which may also effect the rate of dissociation and thus preference of removal of these subunits. [22]

Correlation between experimentally observed and predicted percolation thresholds indicates the applicability of percoalation theory to the study of capsid assembly and disassembly. During HBV disassembly, species missing more than 30 (~25%) subunits appear likely to break into multiple fragments (this work and reference [23]). This result is immediately applicable to assembly reactions. During assembly the intermediate reactions are generally modeled as reversible [2,4,13,54,55]; this implies that dissociation of a subunit from HBV complexes of ≤ 90 dimers may lead to a catastrophic collapse. As icosahedral capsids can be described in terms of relatively few classes of lattices, we suggest that there may be few and generalizable disassembly models, at least some of which will resemble our results based on HBV [56,57]. Furthermore, the implications of these studies relate to the biology and nanotechnology of capsid-like particles uncoating to release cargo, which may include genetic material and therapeutic small molecules [58,59].

Acknowledgments

This work was supported by NIH R01-AI118933 to AZ.

References

1. Schrenk KJ, Hilário MR, Sidoravicius V, Araújo NAM, Herrmann HJ, et al. Critical Fragmentation Properties of Random Drilling: How Many Holes Need to Be Drilled to Collapse a Wooden Cube? Phys Rev Lett. 2016; 116:055701. [PubMed: 26894717]
2. Perlmutter JD, Hagan MF. Mechanisms of virus assembly. Annu Rev Phys Chem. 2015; 66:217–239. [PubMed: 25532951]
3. Ceres P, Zlotnick A. Weak protein-protein interactions are sufficient to drive assembly of hepatitis B virus capsids. Biochemistry. 2002; 41:11525–11531. [PubMed: 12269796]
4. Zlotnick A. To build a virus capsid. An equilibrium model of the self assembly of polyhedral protein complexes. J Mol Biol. 1994; 241:59–67. [PubMed: 8051707]
5. Johnson JM, Tang J, Nyame Y, Willits D, Young MJ, et al. Regulating self-assembly of spherical oligomers. Nano Lett. 2005; 5:765–770. [PubMed: 15826125]
6. Tresset G, Le Coeur C, Bryche JF, Tatou M, Zeghal M, et al. Norovirus capsid proteins self-assemble through biphasic kinetics via long-lived stave-like intermediates. Journal of the American Chemical Society. 2013; 135:15373–15381. [PubMed: 23822934]
7. Zlotnick A, Johnson JM, Wingfield PW, Stahl SJ, Endres D. A theoretical model successfully identifies features of hepatitis B virus capsid assembly. Biochemistry. 1999; 38:14644–14652. [PubMed: 10545189]
8. Zandi R, van der Schoot P, Reguera D, Kegel W, Reiss H. Classical nucleation theory of virus capsids. Biophys J. 2006; 90:1939–1948. [PubMed: 16387781]
9. Packianathan C, Katen SP, Dann CE 3rd, Zlotnick A. Conformational changes in the Hepatitis B virus core protein are consistent with a role for allostery in virus assembly. J Virol. 2010; 84:1607–1615. [PubMed: 19939922]
10. Lazaro GR, Hagan MF. Allosteric Control of Icosahedral Capsid Assembly. J Phys Chem B. 2016; 120:6306–6318. [PubMed: 27117092]
11. Hicks SD, Henley CL. An irreversible growth model for virus capsid assembly. Phys Rev E Stat Nonlin Soft Matter Phys. 2006; 74:031912. [PubMed: 17025672]
12. Zlotnick A. Distinguishing reversible from irreversible virus capsid assembly. J Mol Biol. 2007; 366:14–18. [PubMed: 17157314]
13. Rapaport DC. Studies of reversible capsid shell growth. Journal of physics Condensed matter : an Institute of Physics journal. 2010; 22:104115. [PubMed: 21389449]
14. Seeger, C., Zoulim, F., Mason, WS. Hepadnaviruses. In: Knipe, DM.Griffin, DE.Lamb, RA.Martin, MA.Roizman, B., et al., editors. Fields Virology. Philadelphia: Lippincott Williams & Wilkins; 2007. p. 2977-3029.
15. Selzer, L., Zlotnick, A. Assembly and Release of Hepatitis B Virus. In: Seeger, C., Locarnini, S., editors. Hepatitis B and Delta Viruses. New York: Cold Spring Harbor Press; 2014.
16. Caspar DLD, Klug A. Physical principles in the construction of regular viruses. Cold Spring Harbor Symp Quant Biol. 1962; 27:1–24. [PubMed: 14019094]
17. Uetrecht C, Watts NR, Stahl SJ, Wingfield PT, Steven AC, et al. Subunit exchange rates in Hepatitis B virus capsids are geometry- and temperature-dependent. Phys Chem Chem Phys. 2010; 12:13368–13371. [PubMed: 20676421]
18. Singh S, Zlotnick A. Observed hysteresis of virus capsid disassembly is implicit in kinetic models of assembly. J Biol Chem. 2003; 278:18249–18255. [PubMed: 12639968]
19. Wu G, Liu B, Zhang Y, Li J, Arzumanyan A, et al. Preclinical Characterization of GLS4, an Inhibitor of Hepatitis B Virus Core Particle Assembly. Antimicrobial agents and chemotherapy. 2013; 57:5344–5354. [PubMed: 23959305]
20. Zlotnick A, Venkatakrishnan B, Tan Z, Lewellyn EB, Turner WW, et al. Core protein: a pleiotropic keystone in the HBV lifecycle. Antiviral Research. 2015; 181:82–93.
21. Venkatakrishnan B, Katen SP, Francis S, Chirapu S, Finn MG, et al. Hepatitis B Virus Capsids Have Diverse Structural Responses to Small-Molecule Ligands Bound to the Heteroaryldihydropyrimidine Pocket. J Virol. 2016; 90:3994–4004. [PubMed: 26842475]

22. Venkatakrishnan B, Zlotnick A. The Structural Biology of Hepatitis B Virus: Form and Function. Annual Reviews of Virology. 2016; 3:429–451.
23. Lee LS, Brunk N, Haywood DG, Keifer DZ, Pierson EE, et al. A Molecular Breadboard: Removal and Replacement of Subunits in an Hepatitis B Virus Capsid Protein Science. 2017 epub Aug 10, 2017.
24. Pierson EE, Keifer DZ, Selzer L, Lee LS, Contino NC, et al. Detection of late intermediates in virus capsid assembly by charge detection mass spectrometry. Journal of the American Chemical Society. 2014; 136:3536–3541. [PubMed: 24548133]
25. Harms ZD, Selzer L, Zlotnick A, Jacobson SC. Monitoring Assembly of Virus Capsids with Nanofluidic Devices. ACS Nano. 2015; 9:9087–9096. [PubMed: 26266555]
26. Kondylis P, Zhou Z, Harms ZD, Kneller AR, Lee LS, et al. Nanofluidic Devices with 8 Pores in Series for Real-Time, Resistive-Pulse Analysis of Hepatitis B Virus Capsid Assembly. Anal Chem. 2017; 89:4855–4862. [PubMed: 28322548]
27. Golden KM, Ackley SF, Lytle VV. The percolation phase transition in sea Ice. Science. 1998; 282:2238–2241. [PubMed: 9856942]
28. Onsager L. Crystal statistics I A two-dimensional model with an order-disorder transition. Physical Review. 1944; 65:117–149.
29. Jeong H, Tombor B, Albert R, Oltvai ZN, Barabasi AL. The large-scale organization of metabolic networks. Nature. 2000; 407:651–654. [PubMed: 11034217]
30. Callaway DS, Newman MEJ, Strogatz SH, Watts DJ. Network robustness and fragility: Percolation on random graphs. Physical Review Letters. 2000; 85:5468–5471. [PubMed: 11136023]
31. Newman MEJ. The Structure and Function of Complex Networks. SIAM Review. 2003; 45:167–256.
32. Albert R, Barabasi AL. Statistical mechanics of complex networks. Reviews of Modern Physics. 2002; 74:47–97.
33. Bredon G. Topology and geometry. Graduate Texts in Mathematics. 1997; 139
34. Watts DJ, Strogatz SH. Collective dynamics of ‘small-world’ networks. Nature. 1998; 393:440–442. [PubMed: 9623998]
35. Bashan A, Berezin Y, Buldyrev SV, Havlin S. The extreme vulnerability of interdependent spatially embedded networks. Nature Physics. 2013; 9:667–672.
36. Pierson E, Keifer DZ, Kukreja AA, Wang JC-Y, Zlotnick A, et al. Charge Detection Mass Spectrometry Identifies Preferred Non-Icosahedral Polymorphs in the Self-Assembly of Woodchuck Hepatitis Virus Capsids. J Mol Biol. 2016; 428:292–300. [PubMed: 26151485]
37. Lutomski CA, Lyktey NA, Zhao Z, Pierson EE, Zlotnick A, et al. Hepatitis B Virus Capsid Completion Occurs through Error Correction. J Am Chem Soc. 2017; 139:16932–16938. [PubMed: 29125756]
38. Uetrecht C, Barbu IM, Shoemaker GK, van Duijn E, Heck AJ. Interrogating viral capsid assembly with ion mobility-mass spectrometry. Nat Chem. 2011; 3:126–132. [PubMed: 21258385]
39. Zlotnick A, Aldrich R, Johnson JM, Ceres P, Young MJ. Mechanism of capsid assembly for an icosahedral plant virus. Virology. 2000; 277:450–456. [PubMed: 11080492]
40. Ivanovska IL, Miranda R, Carrascosa JL, Wuite GJ, Schmidt CF. Discrete fracture patterns of virus shells reveal mechanical building blocks. Proc Natl Acad Sci U S A. 2011; 108:12611–12616. [PubMed: 21768340]
41. Reguera J, Carreira A, Riolobos L, Almendral JM, Mateu MG. Role of interfacial amino acid residues in assembly, stability, and conformation of a spherical virus capsid. Proc Natl Acad Sci U S A. 2004; 101:2724–2729. [PubMed: 14981262]
42. Boyd KJ, Bansal P, Feng J, May ER. Stability of Norwalk Virus Capsid Protein Interfaces Evaluated by in Silico Nanoindentation. Front Bioeng Biotechnol. 2015; 3:103. [PubMed: 26284238]
43. Polles G, Indelicato G, Potestio R, Cermelli P, Twarock R, et al. Mechanical and assembly units of viral capsids identified via quasi-rigid domain decomposition. PLoS Comput Biol. 2013; 9:e1003331. [PubMed: 24244139]

44. Schlicksup CJ, Wang JC, Francis S, Venkatakrishnan B, Turner WW, et al. Hepatitis B virus core protein allosteric modulators can distort and disrupt intact capsids. Elife. 2018; 7
45. Zakalyukin RM, Chizhikov VA. Calculations of the Percolation Thresholds of a Three-Dimensional (Icosahedral) Penrose Tiling by the Cubic Approximant Method. Crystallography Reports. 2005; 50:938–948.
46. Arkhipov A, Roos WH, Wuite GJ, Schulten K. Elucidating the mechanism behind irreversible deformation of viral capsids. Biophysical journal. 2009; 97:2061–2069. [PubMed: 19804738]
47. Roos WH, Ivanovska IL, Evilevitch A, Wuite GJ. Viral capsids: mechanical characteristics, genome packaging and delivery mechanisms. Cellular and molecular life sciences : CMLS. 2007; 64:1484–1497. [PubMed: 17440680]
48. van Rosmalen MG, Roos WH, Wuite GJ. Material properties of viral nanocages explored by atomic force microscopy. Methods Mol Biol. 2015; 1252:115–137. [PubMed: 25358778]
49. Zeng C, Moller-Tank S, Asokan A, Dragnea B. Probing the Link among Genomic Cargo, Contact Mechanics, and Nanoindentation in Recombinant Adeno-Associated Virus 2. J Phys Chem B. 2017; 121:1843–1853. [PubMed: 28142241]
50. Moreno-Madrid F, Martin-Gonzalez N, Llauro A, Ortega-Esteban A, Hernando-Perez M, et al. Atomic force microscopy of virus shells. Biochem Soc Trans. 2017; 45:499–511. [PubMed: 28408490]
51. Caspar DL, Klug A. Physical principles in the construction of regular viruses. Cold Spring Harb Symp Quant Biol. 1962; 27:1–24. [PubMed: 14019094]
52. Oosawa, F., Asakura, S. Thermodynamics of polymerization of protein. London: Academic Press; 1975.
53. Abete T, de Candia A, Lairez D, Coniglio A. Percolation model for enzyme gel degradation. Phys Rev Lett. 2004; 93:228301. [PubMed: 15601123]
54. Porterfield, JZ., Zlotnick, A. An Overview of Capsid Assembly Kinetics. In: Stockley, PG., Twarock, R., editors. Emerging Topics in Physical Virology. London: Imperial College Press; 2010.
55. Hagan MF, Elrad OM. Understanding the concentration dependence of viral capsid assembly kinetics–the origin of the lag time and identifying the critical nucleus size. Biophys J. 2010; 98:1065–1074. [PubMed: 20303864]
56. Mannige RV, Brooks CL 3rd. Tilable nature of virus capsids and the role of topological constraints in natural capsid design. Phys Rev E Stat Nonlin Soft Matter Phys. 2008; 77:051902. [PubMed: 18643097]
57. Nguyen HD, Brooks CL 3rd. Generalized structural polymorphism in self-assembled viral particles. Nano Lett. 2008; 8:4574–4581. [PubMed: 19367856]
58. Czapar AE, Steinmetz NF. Plant viruses and bacteriophages for drug delivery in medicine and biotechnology. Curr Opin Chem Biol. 2017; 38:108–116. [PubMed: 28426952]
59. Madigan VJ, Asokan A. Engineering AAV receptor footprints for gene therapy. Curr Opin Virol. 2016; 18:89–96. [PubMed: 27262111]

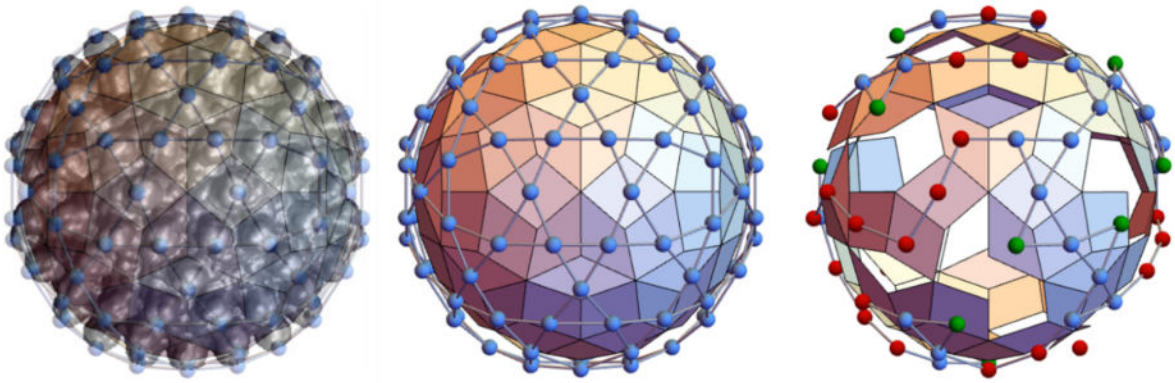

Figure 1. The T = 4 HBV capsid and corresponding graph representations of a fully assembled and partially disassembled capsid

A) A $T = 4$ HBV structure (PDB accession code 2G33) overlaid with its dual graph. The capsid is built of 120 homodimeric protein subunits. The dual vertices are represented as blue spheres. Interstitial lines on the surface of the capsid correspond to the 240 chemical contacts between dimers. The physical diameter of the $T = 4$ HBV capsid is 36 nm. (B) A simpler polyhedral model of the $T = 4$ HBV capsid, using interstitial lines to demarcate subunits, and its dual graph. C) The HBV graph missing $i = 30$ random subunits. The vertex color indicates the local clustering coefficient, $C(v_i)$. Weak, linear-chain subunits have $C(v_i) = 0$ (red). Partially isolated vertices that are members of trimeric clusters have $C(v_i) = 1$ (green). The remaining subunits have an environment that is closer to that of a subunit in an intact particle with $C(v_i) = \frac{1}{3}$(blue). See the Definitions section for the equation determining a vertex's local clustering coefficient.

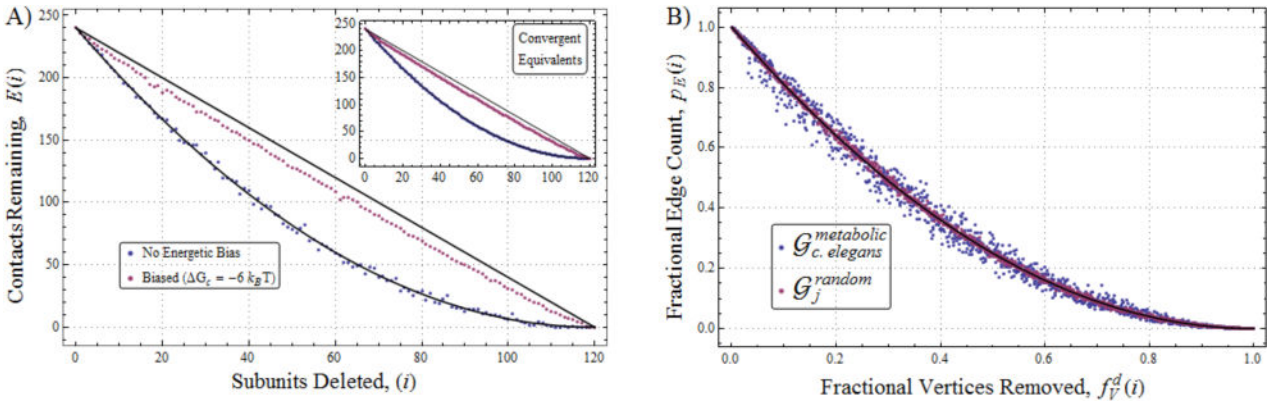

Figure 2. Graph stability decays quadratically as vertices are removed

Stability is proportional to the number of edges between vertices, or equivalently, to the number of inter-subunits contacts remaining after removing *i* subunits. (A) Using the 4-regular HBV graph, data were collected using a single Monte Carlo replicate for each value of *i*. Agreement with the fit to Equation 3.3 (black line) is unambiguous. In a biased reaction interactions between subunits favor expansion of a single hole, resulting in a linear decrease in the number of intersubunit contacts. After several replicates, points converge to the line (inset panel). (B) All tested data graphs follow the same curve, they do not have to be regular. Graphs used in this figure include 1173 Erdos-Renyi random graphs of low but non-zero variance from k=4 (purple), and the highly variant metabolic network of C. elegans (blue) [29]. The goodness of fit to the quadratic curve of Equation 3.3 (black line) is related to the degree distributions and replication; the HBV data fall very close to the calculated curve after 100 replications (inset)

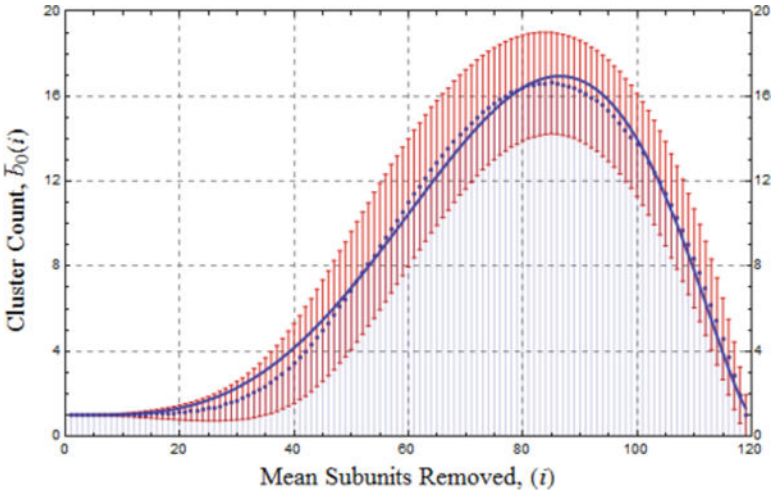

Figure 3. The number of fragmented clusters $\bar{b}_0(\mathbf{i})$ increases abruptly as vertices are removed from the HBV graph

Blue data points are the averages of $b_0(i)$ over 100,000 Monte Carlo replicates for each value of i. The red bars indicate the standard deviations. The extent of fragmentation increases rapidly in the range $30 \leq i \leq 40$. The data for the dashed line were described previously [23]. The solid blue line is approximated well by a β-curve with parameterization $\alpha \approx 3.73$, $\beta \approx 1.45$, and $c_n \approx 343$ (Equation 3.1).

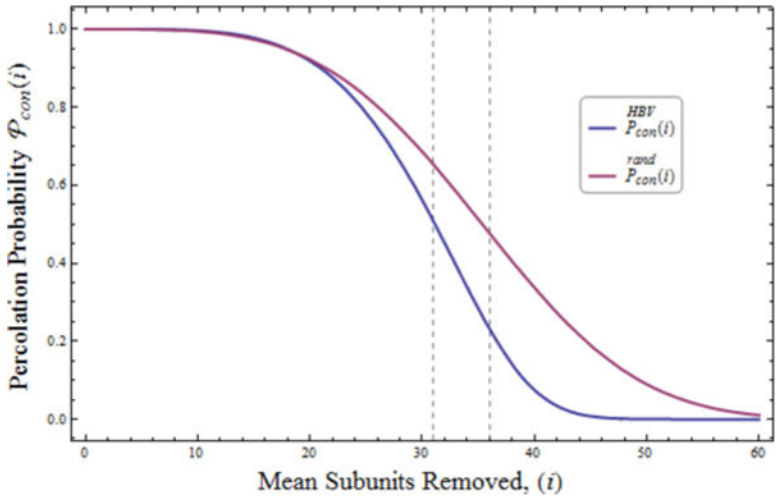

Figure 4. For a graph modeled on the HBV capsid, the percolation threshold occurs more abruptly than for non-planar topologies

Shown is the probability the perforated capsid remains intact $P_{con}(i)$ for both HBV and an array of 120 subunit 4-regular – but otherwise random – graphs. The leftmost vertical gridline at $i_c^{HBV} = 31$ corresponds to the percolation threshold of HBV, p_{V_c}. This threshold has been confirmed *in vitro*: we see few large fragments missing i_c^{HBV} or more subunits. HBV is more fragile than non-planar systems, fragmenting more abruptly than comparable random graphs due to the difference in planar (2D) and non-planar topologies: $i_c^{HBV} < i_c^{rand}$.

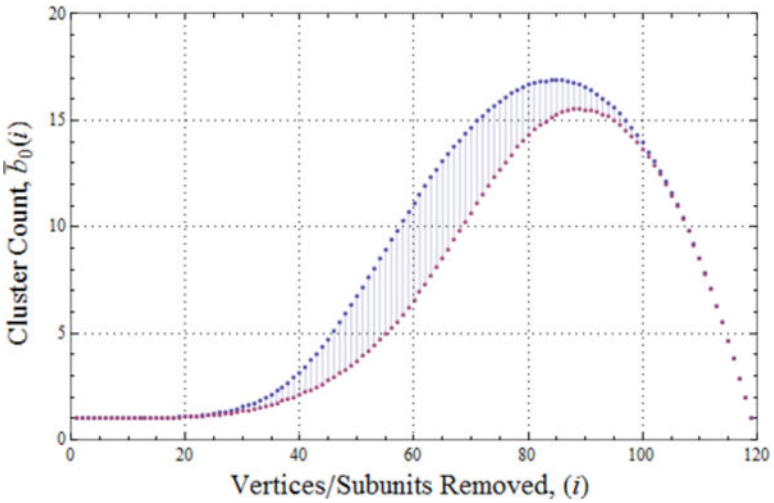

Figure 5. The graph corresponding to HBV capsid, a planar tiling, fragments with fewer subunits removed and into more fragments than non-planar graphs

The number of fragmented clusters $\bar{b}_0(i)$ was determined as a function of the number of vertices removed for the HBV-like graph (blue, data from Figure 3) and an ensemble of 4-regular random graphs (red). The HBV lattice is more susceptible to fragmentation because vertices connect only with nearest neighbors in two dimensions.

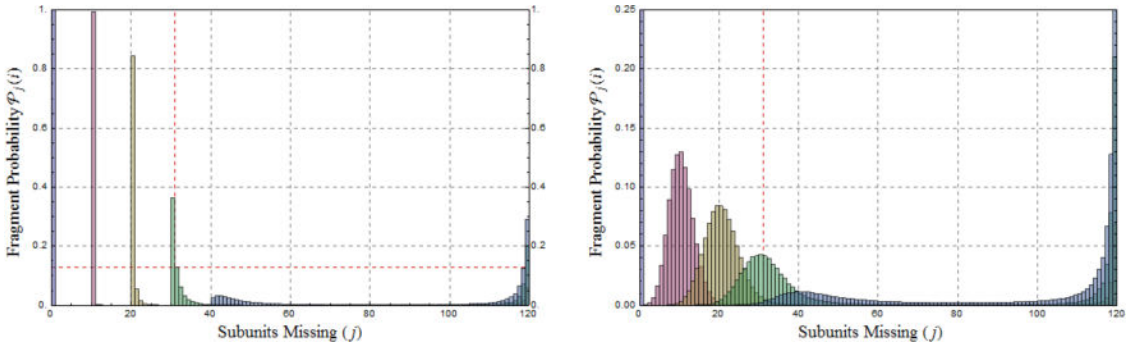

Figure 6. Fragment size distributions $P_j(i)$ at various i

As subunits are removed, fragmentation causes a rapid decrease in the probability of intact, perforated capsids. (A) Monte Carlo disassembly of HBV produces a mean number of fragmented clusters, $\bar{b}_0(i)$, data and standard deviations for which are shown. The probability of intact clusters missing i subunits are shown as histogram bins; color denotes the value of i = (0, 10,…, 40) at which each distribution was produced. At any given deterministic i, the largest cluster possible occurs with probability $P_{con} = P_{120-i}$, the leftmost bin of each color. The probability of one large, intact cluster P_{120-i} remains substantial until reaching the percolation threshold corresponding to $i_c = 31$ subunits removed, at which $P_{90}(30) \approx .35$ or 35%. Note that when $i = 40 > i_c$, there exist as many fully isolated subunits as there are intact wholes when $i = 30 < i_c$: $P_{90}(30) \approx P_1(40)$. We have detected this percolation threshold *in vitro* [23]. Panel A was described previously and is used by permission [23]. (B) The same for a more realistic case in which the number of subunits removed per capsid is normally distributed about a mean i. Such stochasticity further reduces $P_j(i)$, explaining lack of detection of many clusters of size $j < 90$.

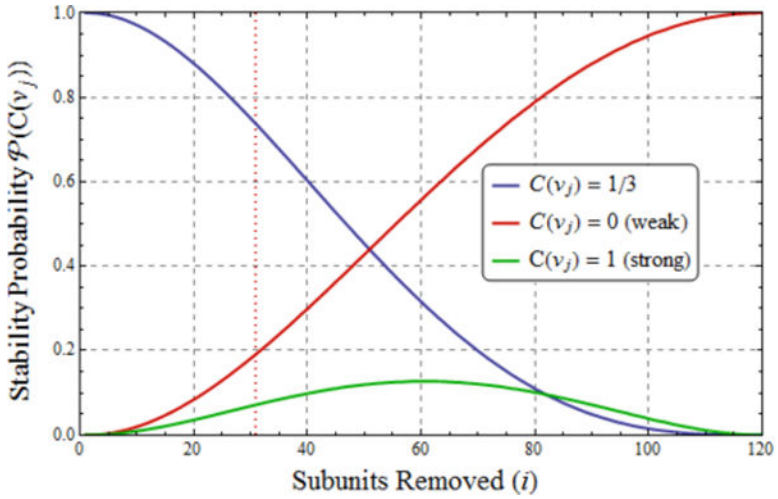

Figure 7. Evolution of the local subunit stability in a dissociation reaction

The clustering coefficient indicates the local integrity and stability of products. The same color scheme is used in the left panel of figure 1. The fraction of capsid-like subunits (blue, $C\left(v_j\right) = \frac{1}{3}$) falls monotonically. The proportion of subunits in easily fragmented, linear chains (red, $C(v_j) = 0$) increases rapidly. It is likely these will further dissociate in solution. The fraction of subunits in small, trimeric clusters (green, $C(v_j) = 1$) is maximal at about the point when half of the subunits are removed. The probability of even small, trimeric clusters begins decreasing after $i \approx 60$ in favor of completely isolated subunits.

Table 1

A glossary of variables used in this paper.

Symbol	Relation	Definition
G		The initial graph (prior to disassembly)
V_0		Initial number of vertices (prior to disassembly)
E_0		Initial number of edges
$k_0, \bar{k}_0$	$= \left(\frac{E_0}{V_0}\right)$	Initial valence (number of edges per vertex). The bar indicates the mean valence.
$G(i)$		A graph missing (i) vertices
$V(i)$	$= V_0 - i$	Number of vertices remaining in $G(i)$
$p_V(i)$	$= \left(\frac{V(i)}{V_0}\right)$	Fraction of vertices remaining in $G(i)$
$f_V^d(i)$	$= \left(\frac{i}{V_0}\right)$ $= 1 - \rho_V(i)$	Fraction of vertices removed from G to produce $G(i)$
$E(i)$ $\bar{E}(i)$		Number of edges remaining in $G(i)$. The bar indicates the mean of multiple Monte Carlo replicates.
$p_E(i)$	$= \left(\frac{E(i)}{E_0}\right)$	Fraction of edges remaining in $G(i)$
$b_0(0)$		Initial number of components
$b_0(i)$ $\bar{b}_0(i)$		Number of fragments resulting from removing i vertices to make $G(i)$. The bar indicates the mean of multiple Monte Carlo replicates.
a, β, c_n		Fit parameters characterizing the tendency to fragment.
$\bar{P}_j(i)$		Mean probability that a fragment of size (j) exists in a given $G(i)$
i_c		Percolation threshold, number of vertices removed to induce a disconnection phase transition.
p_{V_c}	$= \left(\frac{V_0 - i_c}{V_0}\right)$	Percolation threshold in terms of fractional vertices remaining.
$C(v_j)$		Clustering index indicates the class of local connectivity for a vertex. Classes, not the value of $C(v_j)$, are indicative of local stability. $C(v_j) = 0$ − locally unstable $C(v_j) = 1$ − locally stable $C(v_j) = 1/3$ − capsid-like environment, stable