Pathways for Virus Assembly around Nucleic Acids

Jason D. Perlmutter, Matthew R. Perkett and Michael F. Hagan

Martin Fisher School of Physics, *Brandeis University, Waltham, MA 02454, USA*

Correspondence to Michael F. Hagan: *hagan@brandeis.edu*
http://dx.doi.org/10.1016/j.jmb.2014.07.004
Edited by C. L. Brooks III

Abstract

Understanding the pathways by which viral capsid proteins assemble around their genomes could identify key intermediates as potential drug targets. In this work, we use computer simulations to characterize assembly over a wide range of capsid protein–protein interaction strengths and solution ionic strengths. We find that assembly pathways can be categorized into two classes, in which intermediates are either predominantly ordered or disordered. Our results suggest that estimating the protein–protein and the protein–genome binding affinities may be sufficient to predict which pathway occurs. Furthermore, the calculated phase diagrams suggest that knowledge of the dominant assembly pathway and its relationship to control parameters could identify optimal strategies to thwart or redirect assembly to block infection. Finally, analysis of simulation trajectories suggests that the two classes of assembly pathways can be distinguished in single-molecule fluorescence correlation spectroscopy or bulk time-resolved small-angle X-ray scattering experiments.

Introduction

In many virus families, the spontaneous assembly of a protein shell (capsid) around the viral nucleic acid (NA) genome is an essential step in the viral life cycle [1]. These families include most viruses with single-stranded RNA (ssRNA) genomes, as well as the Hepadnaviridae (e.g., hepatitis B virus, HBV). Understanding the mechanisms that underlie this cooperative assembly process could facilitate efforts to develop antiviral drugs that block or derail the formation of infectious particles (for reviews, see Refs. [2] and [3]) and promote efforts to reengineer them for biomedical delivery. In this article, we explore how the interactions between the molecular components determine the mechanism of assembly and how these interactions can be altered by changing solution conditions or mutagenesis to modulate assembly pathways.

The most detailed knowledge of capsid–NA interactions comes from structural analysis of assembled viral particles. Atomic-resolution structures of capsids assembled around ssRNA have been obtained by X-ray crystallography and/or cryo-electron microscopy (e.g., Refs. [4–16]). The packaged NAs are less ordered than their protein containers and thus have been more difficult to characterize. However, cryo-electron microscopy experiments have identified that the nucleotide densities are nonuniform, with a peak near the inner capsid surface and relatively low densities in the interior [7,17,18]. While atomistic detail has not been possible in these experiments, all-atom models have been derived from equilibrium simulations [19–21]. In some cases, striking image reconstructions reveal that the packaged RNA adopts the symmetry of the overlying capsid (e.g., Refs. [7,10,16,22], and [23]). While it has been proposed that this order arises as a function of the assembly mechanism for several viruses [24–26], computational analysis of polyelectrolyte configurations inside capsids also indicate that capsid–polymer interactions can generically drive spatial organization of the packaged polymer [20,27–38]. Theoretical works have also characterized the relationship between the NA charge and structure and the length that is optimal for packaging [27,31,32,38–45].

In addition to this structural data on assembled capsids, an extensive combination of mass spectrometry, assembly kinetics experiments, constraints from

assembled capsid structures, and mathematical modeling has delineated assembly pathways for several viruses, with a particular focus on the role of interactions between capsid proteins and specific RNA sequences called "packaging signals". Recent single-molecule fluorescence correlation spectroscopy (smFCS) experiments indicate that, for these viruses, assembly around the viral genome is more robust and proceeds by a different mechanism as compared to around heterologous RNA [46]. However, in other cases, capsid proteins show no preference for genomic RNA over heterologous RNA (e.g., HBV [47]), and cowpea chlorotic mottle virus (CCMV) proteins preferentially encapsidate heterologous RNA [from Brome mosaic virus (BMV)] over the genomic CCMV RNA with equivalent length [48]. Furthermore, experimental model systems in which capsid proteins assemble into icosahedral capsids around synthetic polyelectrolytes or other polyanions [49–61] demonstrate that specific RNA sequences are not required for capsid formation or cargo packaging. Thus, a complete picture of capsid assembly mechanisms requires understanding how assembly pathways depend on those features that are generic to polyelectrolytes, as well as those which are specific to viral RNAs.

In previous work on assembly around a simple model for a polymer, Elrad and Hagan proposed that mechanisms for assembly around a cargo (i.e., RNA, polymer, or nanoparticle) can be classified on the basis of two extreme limits [37]. In the first (originally proposed by McPherson [62] and then by Hagan [63] and Devkota *et al.* [20]), strong protein–cargo interactions drive proteins to adsorb "en masse" onto the cargo in a disordered manner, meaning that there are few protein–protein interactions. Once enough subunits are bound, subunits undergo cooperative rearrangements (potentially including dissociation of excess subunits) to form an ordered capsid. This mechanism has been observed in recent simulations [37,38,63–65]. In the second limit, where protein–protein interactions dominate, a small partial capsid nucleates on the cargo, followed by a growth phase in which individual proteins or small oligomers sequentially add to the growing capsid. This class of pathways resembles the nucleation-and-growth mechanism by which empty capsids assemble [66], except that the polymer plays an active role by stabilizing protein–protein interactions and by enhancing the flux of proteins to the assembling capsid [37,67,68].

It is difficult to determine assembly mechanisms directly from experiments due to the small size ($\lesssim$10 of nm) and transience (~ms) of most intermediates. Observations *in vitro* suggest that both mechanisms may be viable. Kler *et al.* used time-resolved small-angle X-ray scattering (trSAXS) to monitor simian virus 40 (SV40) capsid proteins assembling around ssRNA [69]. The scattering profiles at all time points during assembly could be decomposed into unassembled components (RNA + protein subunits) and complete capsid; the absence of any signal corresponding to a large disordered intermediate suggests that this assembly follows the nucleation-and-growth (ordered) assembly mechanism [69]. Other observations suggest that viruses can assemble through the *en masse* mechanism. Garmann *et al*. and Cadena-Nava *et al*. found that *in vitro* assembly of CCMV was most productive when performed in two steps [70,71]: (1) at low salt (strong protein–RNA interactions) and neutral pH (weak protein–protein interactions), the proteins undergo extensive adsorption onto RNA, then (2) pH is reduced to activate binding of protein–protein binding [70]. Similarly, a recent observation of capsid protein assembly around charge-functionalized nanoparticles found that assembly initially proceeded through nonspecific aggregation of proteins and nanoparticles, followed by the gradual extrusion of nanoparticles within completed capsids [72]. These experiments used viral proteins with relatively weak protein–protein interactions (CCMV and BMV) [73] and moderate salt concentrations (100–150 mM). The experiments of Kler *et al.* considered SV40 proteins [69,74], which have strong protein–protein interactions [73] and high salt (250 mM). Together, these *in vitro* experiments suggest that productive assembly could proceed by either the *en masse* or the nucleation-and-growth mechanism.

In this work, we use dynamical simulations to investigate the extent to which the assembly mechanism can be controlled by tuning solution ionic strength and protein–protein attractions. We extend a model that was recently used to calculate the thermostability and assembly yields of viral particles as a function of protein charge and NA length and structure. Those previous simulations found quantitative agreement between predicted NA lengths that optimize capsid thermostability and viral genome length for seven viruses [38]. Here, we perform extensive new simulations of assembly, in which protein–protein interactions, the sequence of charges in capsid protein–NA binding domains, and the solution ionic strength are varied. We find that, by varying these control parameters, the assembly mechanism can be systematically varied between the two extreme limits described above. Our results suggest that knowledge of protein–protein and protein–NA binding affinities may be sufficient to predict which assembly mechanism will occur, and we estimate relative protein–NA binding interactions for three viruses (based on nonspecific electrostatic interactions). These findings suggest that assembly mechanisms can be rationally designed through choice of solution conditions and mutagenesis of capsid protein–protein interfaces and protein–NA binding domains. Finally, by calculating hydrodynamic radii and small-angle X-ray scattering (SAXS) profiles associated with assembly intermediates, we