Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Virology

Recent advances in coarse-grained modeling of virus assembly

Michael F Hagan[1] and Roya Zandi[2]

Addresses
[1] Martin Fisher School of Physics, Brandeis University, Waltham, MA 02453, USA
[2] Department of Physics and Astronomy, University of California, Riverside, California 92521, USA

Corresponding authors: Hagan, Michael F (hagan@brandeis.edu) and Zandi, Roya (royaz@ucr.edu)

Current Opinion in Virology 2016, **18**:36–43

This review comes from a themed issue on **Virus structure and expression**

Edited by **Adam Zlotnik** and **Juha Huiskonen**

http://dx.doi.org/10.1016/j.coviro.2016.02.012

1879-6257/Published by Elsevier B.V.

Capsid assembly and packaging of the genome are essential steps in the formation of an infectious virus. Thus, elucidating the mechanisms by which assembly proceeds could identify important targets for antiviral drugs and would advance our fundamental understanding of the viral lifecycle. However, assembly and packaging pathways remain incompletely understood for many viruses because most intermediates are transient and therefore undetectable or characterized only with low resolution. Computer simulations of virus assembly have overcome this limitation by revealing features of assembly processes that are not accessible to experiments alone. However, the large size of a virus (15–1000 nm) and the timescales required for assembly (ms-hours) prohibit simulating capsid formation with atomic-resolution, except for specific steps [1]. To this end, researchers have relied on simplified models, which aim to coarse-grain over atomic-scale details while accurately describing the essential physical features that control assembly.

This review describes recent advances in coarse-grained models of capsid assembly. We begin with a brief overview of models and simulation methodologies, followed by recent applications of these approaches. To accommodate space limitations, we limit our discussion of applications to two areas which have recently been the subject of intense modeling activity: the role of nucleic acids in the assembly of icosahedral viruses, and assembly of the mature HIV capsid.

Coarse-grained models for capsid assembly

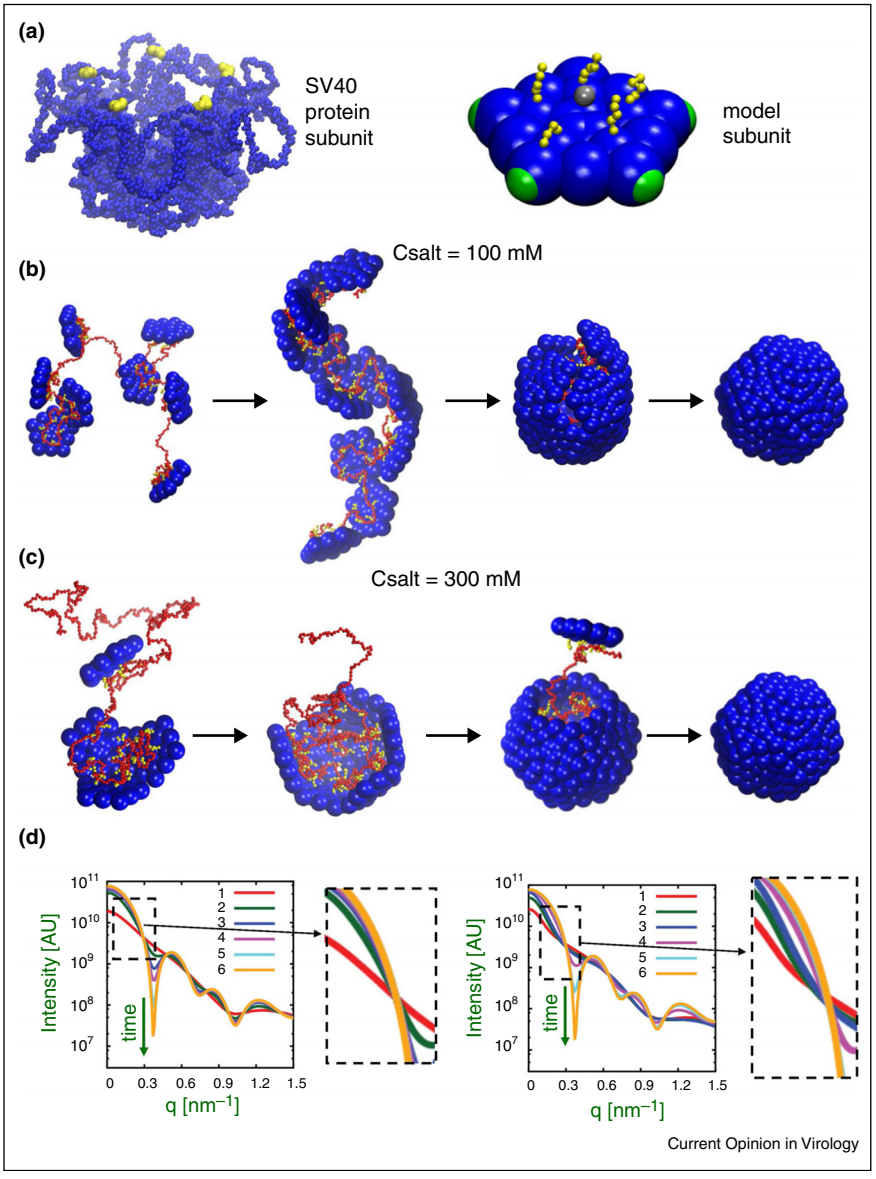
One approach to model development seeks to describe a specific physical system with the greatest accuracy allowed by computational constraints, and by systematically coarse-graining from atomistic simulations (e.g. [2,3]). However, the conformational dynamics of capsid proteins restricts the accuracy of such techniques, and the complexity of the resulting coarse-grained models has limited their application to assembly. Therefore, capsid assembly models have relied on a combination of atomistic simulations, structural data, and fitting model parameters to kinetics and thermodynamic data. Often, the aim has been to construct the simplest model consistent with experimental data, to discover general, fundamental insights about capsid assembly.

Models for virus assembly can be separated into three classes. In the first, the time evolution of concentrations of capsid intermediates is represented by a system of rate equations [4,5•]. Formulation of the model requires specifying the state space (the set of all possible assembly intermediates) and transition rates between each pair of intermediates. The rate equations can be numerically integrated [4,5•,6–8], or trajectories consistent with the rate equations can be stochastically sampled using Gillespie-type algorithms [9–13,14••], and transition rates can be fit against experimental data [4,12,15,16]. Despite their simplicity, such models reproduce many experimental observations on capsid assembly.

In the next class of models (particle-based simulations), subunits interact through pair potentials that drive assembly toward an ordered low-energy structure (e.g. an icosahedral shell [17–22,23•,24,25], Figure 1). Subunit motions are explicitly tracked by numerically integrating equations of motion (e.g. using molecular dynamics, Brownian dynamics (BD), or discontinuous molecular dynamics [19,21,23•]. The third approach combines aspects of Gillespie-type and particle-based models, modeling assembly through the irreversible addition of triangular subunits to growing edges of an incomplete shell [26–28]. The shell is treated as an elastic sheet that relaxes to its minimum energy configuration after each accretion.

Of these approaches, particle-based simulations enable (at least in principle) the fewest assumptions about intermediate geometries and the highest resolution description of proteins. However, their high computational cost has limited model resolution, simulation timescales, and system sizes [29,30].

Figure 1



A particle-based model for capsid assembly around a linear polymer. **(a)** (Left) Image of the crystal structure of a homopentamer of the SV40 capsid protein, the elemental subunit for SV40 capsid assembly, with visible portions of RNA binding domains in yellow. (Right) Image of a coarse-grained model subunit, with RNA binding domains shown in yellow. **(b,c)** Snapshots from typical simulation trajectories illustrating two classes of pathways for assembly around a polymer or RNA. In (b), strong protein–RNA interactions lead to an 'en masse' mechanism, in which proteins rapidly adsorb onto the RNA in a disordered manner, followed by cooperative rearrangements to form an assembled capsid. In (c), weaker protein-RNA interactions drive a nucleation-and-growth mechanism, in which a small, ordered nucleus forms, followed by sequential addition of protein subunits. **(d)** Time-resolved small angle X-ray scattering (SAXS) profiles estimated from simulation trajectories corresponding to the (left) nucleation-and-growth mechanism and (right) en masse mechanism. The zoom-in to the right of each plot illustrates one of the distinguishing features of the SAXS profiles — the nucleation-and-growth mechanism leads to an isosbestic point among profiles measured at different times, whereas the en masse mechanism does not (at early times).
Figure adapted from Ref [23•].

Recent algorithmic advances are beginning to overcome this limitation. For example, Grime and Voth [31•] designed an efficient parallelization scheme for spatially heterogeneous particle concentrations (which occur during assembly simulations with implicit solvent). Algorithms performing rigid body dynamics on GPUs show significant speedup in comparison to conventional CPUs. For example, the package HOOMD [32] has been used to simulate virus assembly around nucleic acids and on membranes [23•,24,33••,34]. Zuckerman