Inferring Fine-Scale Kinetics of *in vitro* Capsid Assembly via Multi-Curve Fitting

Lu Xie[1], Gregory R. Smith[2], Xian Feng[2], and Russell Schwartz[2, 3]

***Abstract* — Predictive biological simulations depend on accurately physical parameters that we often cannot directly measure. This issue is a particular problem for self-assembly, a ubiquitous phenomenon in cellular systems, where we often lack the ability to experimentally probe individual steps in a complicated cascade of reactions. Here, we address the parameter inference problem in the context of virus capsid assembly, an important model system of macromolecular self-assembly, by using simulation-based data fitting to match parameters of stochastic simulations to indirect measures of bulk capsid assembly *in vitro*. We apply our method to three viral systems: human papillomavirus (HPV), cowpea chlorotic mottle virus (CCMV) and hepatitis B virus (HBV). The results suggest a surprising diversity and complexity of assembly methods for the three systems. Such methods may have much broader value for instantiating models of complex biological systems for which we lack direct experimental methods.**

***Keywords* — Self-Assembly, Stochastic Simulation, Optimization, Model Fitting, Virology.**

I. Motivation

VIRUS capsid assembly is a key model system for complicated self-assembly processes, which has attracted considerable interest across different modeling communities[1]. Simulation methods have proven valuable for characterizing possible mechanisms of capsid assembly, understanding potential pitfalls to efficient assembly, and learning strategies by which viruses overcome them. Simulation studies so far, however, have been able to say little about the assembly kinetics of any specific virus, in large part because we lack detailed reaction rate parameters needed to parameterize simulations. No current experimental method can measure rates of individual steps in the assembly reactions and the limited kinetic data available provides only bulk averages over the assembly of many capsids *in vitro* from purified viral coat proteins [2-4].

To address this problem, we previously developed a simulation-based data fitting method to learn rate parameters consistent with both structure-based rule sets and experimental light scattering data and applied it to bulk assembly progress of HPV *in vitro* [5]. In the present work, we extend the previous method to fit multiple time courses simultaneously in order to reduce possible ambiguity in parameters. We further apply the work to a collection of three icosahedral viruses – HPV [2], CCMV [3], and HBV [4] –to test the generality of the methods and gain insight into the diversity of mechanisms used by different real viruses.

Acknowledgements: This work was supported by NIH grant 1R01AI076318.
[1]Joint Carnegie Mellon – University of Pittsburgh Ph.D Program in Computational Biology and [2]Department of Biological Sciences and [3]Ray and Stephanie Lane Center for Computational Biology, Carnegie Mellon University.

II. Modeling and Methods

Our methods rely on stochastic simulations of capsid assembly developed in our prior work [6], which implement a coarse-grained model of capsid assembly using interaction rules describing binding sites, binding preferences, and interaction rates for interactions of basic assembly subunits (coat proteins or small oligomers). We fit rate parameters to light scattering measurements of bulk *in vitro* assembly by minimizing a root mean squared deviation (RMSD) between the experimental curves and artificial curves implied by the simulations. Our method implements gradient-based and response-surface local optimization methods with a heuristic global search to find an optimal parameter fit. To reduce the potential for multiple local optima, we perform this search simultaneously for multiple light scattering measurements representing distinct assembly concentrations.

III. Results

We apply our method to the three viral systems, learning for each a set of on- and off-rates providing a closest match to the observed data. The resulting fits suggest three very different in vitro assembly mechanisms, with HPV capsid assembly by addition of individual subunits in a non-nucleation limited pathway while the other two viruses show clear features of nucleation limited assembly and more complicated ensembles of assembly pathways involving interactions of distinct sets of oligomer intermediates. These fits provide a basis for exploring how these mechanisms might vary between *in vitro* and *in vivo* assembly conditions.

References

[1] Whitesides GM, Grzybowski B (2002) Self-assembly at all scales. *Science* **295**, 2418-2421.

[2] Casini GL, et al. (2004) *In vitro* papillomavirus capsid assembly analyzed by light scattering. *Virology* **325**, 320-327.

[3] Zlotnick A, et al. (2000) Mechanism of capsid assembly for an icosahedral plant virus. *Virology* **277**, 450-456.

[4] Zlotnick A, et al. (1999) A theoretic model successfully identifies features of Hepatitis B virus capsid assembly. *Biochemistry* **38**, 14644-14652.

[5] Kumar MS, Schwartz R (2010) A parameter estimation technique for stochastic self-assembly systems and its application to human papillomavirus self-assembly. *Phys. Biol.* **7**, 45005-45016.

[6] Zhang T, Rohlfs R, Schwartz R (2005) Implementation of a discrete event simulator for biological self-assembly systems. *Proc. 37th Winter Simulation Conf.* pp. 2223-2231