

THE COMPUTATIONAL MICROSCOPE

Allocation: NSF/18 Mnh, NSF RAPID/0.55 Mnh, Blue Waters Professor/0.24 Mnh
PI: Klaus Schulten¹
Co-PIs: James C. Phillips¹ and John E. Stone¹
Collaborators: Rebecca C. Craven², Costas Maranas³, Eva Nogales^{5,6}, Peijun Zhang⁴

¹University of Illinois at Urbana-Champaign
²Penn State University College of Medicine
³Penn State University
⁴University of Pittsburgh
⁵University of California, Berkeley
⁶Howard Hughes Medical Institute

EXECUTIVE SUMMARY:

FIGURE A: The first computational simulation of the immature lattice of the Rous sarcoma virus. This model provides a template for studying the maturation process of retroviruses as well as evaluating opportunities for drug targets.

Living organisms arise from basic, inanimate matter, but do they blindly following the laws of physics? Computational biology in the petascale era offers a unique opportunity to connect atomic-level descriptions of biological systems (representing inanimate matter) with cell-scale architecture and behavior (representing animate matter). Molecular dynamics (MD) simulations can be used as the “computational microscope” to establish amazing cellular structures at the atomic level, as well as to employ those structures to elucidate the dynamics and underlying physical and chemical mechanisms of cellular processes. Such a computational approach has enjoyed a rapid acknowledgment of its complementary role to experimental observation, the latter often lacking the spatial and temporal resolution needed to truly grasp cellular processes. Here we discuss successful research on several large-scale biological systems.

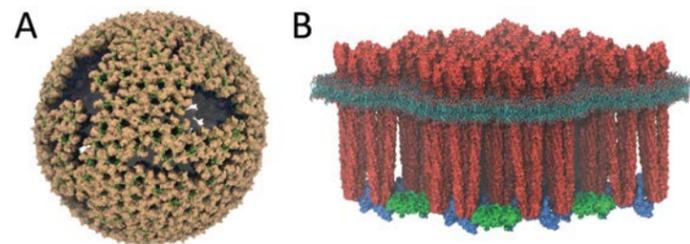
INTRODUCTION

Recent developments on hybrid experimental methods based on the revolutionary advance of electron microscopy have led to previously unimaginable information on cell-level structures that computational modeling requires for solidly based descriptions. Very fortunately, computational modeling can play a significant role in hybrid method structure analysis. First, the accuracy of computational modeling has drastically increased, such that results from computational studies today often exhibit astounding agreement with observation where available. Second, a broad range of so-called sampling methods based on statistical mechanical concepts are developed, enabling the biologically functional timescale in living cells to be covered by these simulations. The projects in this report leverage computational methods to combine structural data from multiple sources of differing resolutions (X-ray crystallography of individual proteins, medium-resolution cryo-EM of multi-protein systems, and low-resolution cryo-EM tomography of subcellular organelles), yielding atomic-resolution structural models of structures on the order of up to 100 nm in size.

METHODS & RESULTS

Since solving the first atomic-level structure of the mature HIV capsid [1], we can now describe the action of host cell factors in capsid stabilization and assembly (cyclophilin A) as well as characterize the effect of a cell factor essential for disrupting the capsid and assisting in nuclear import (TRIM-family proteins). Such studies may help scientists to better understand how the HIV capsid infects the host cells and could lead to new HIV therapies.

While the mature HIV capsid is made only of capsid proteins, the immature lattice requires Gag proteins, a polyprotein essential for its assembly. To investigate the immature virus in atomistic detail, we produced the first all-atom model of the immature lattice of a retrovirus (Figure A) that is closely related to HIV, namely Rous sarcoma virus (RSV) [2]. The atomic structure of the immature RSV lattice represents a milestone for computational modeling, and the structure provides a template to study the maturation process of retroviruses as well as to evaluate opportunities for drug targets.



The recent Ebola outbreak in West Africa prompted us to develop a diagnostic tool that would reliably detect the virus in pre-symptomatic patients. To detect Ebola, we have computationally designed and optimized the first set of antibody-like prototypes. The experimental validation of these prototypes is projected to be completed by early 2016.

Bacteria use large, highly ordered clusters of sensory proteins known as chemosensory arrays (Figure B) to detect and respond to chemicals in their environment. We have integrated multi-scale structural data from experimental sources to computationally construct the first atomic model of the chemosensory array’s molecular architecture. In addition, we identified a novel conformational change in a key signaling protein that is linked at the cellular level to the chemotaxis function (that is, movement in response to chemical stimulus). This model may inspire and assist future experimental and computational studies in elucidating a general mechanistic description of signal transduction in the biological sensory apparatus.

Microtubules are a major component of the cell cytoskeleton, important for maintaining cell structure, intracellular transport, and cell division. Combining multi-scale structural data allowed us to build atomic models for microtubules (Figure C) in different nucleotide binding states (crucial for the switch between phases of assembly and disassembly). Our simulations suggested important structural dynamics events toward the microtubules assemble and stability. Such simulations pave the way to understand the atomic details behind the assembly and disassembly phases of the microtubules as well as the effect of anticancer drugs on microtubule dynamic instability.

Molecular motors, which travel on microtubules, are key drug targets for anti-cancer therapy. Combining novel sampling techniques with the computing power of Blue Waters, we were able to study a long timescale mechanical process (millisecond) driven by chemical energy of a ring-shaped molecular motor [3]. Such a study provides an effective approach to identify possible drug targets on protein structure.

WHY BLUE WATERS

Without Blue Waters and other petascale computing resources, projects involving large molecular systems like HIV, RSV, the chemosensory array, and microtubules would not be possible. These molecular systems are composed by tens of millions of atoms and must be simulated for long periods of time (microseconds). On the other hand, medium systems like the motor protein project (hundreds of thousands of atoms) could be simulated for a timescale of milliseconds only with the computational power delivered by Blue Waters. These projects are examples of how Blue Waters enables bold, new projects that push the limits of what can be done with scientific computing. In our case, that means expanding molecular dynamics simulation capabilities from simulating just a few proteins to simulating full organelles.

FIGURE B: Bacteria chemosensory array model at atomic resolution.

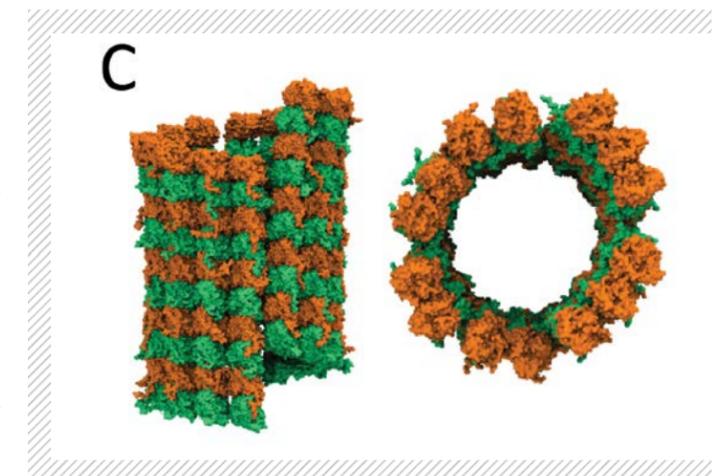


FIGURE C: Side and top view of the atomic model of a microtubule, a multi-protein structure important for cell structure, intracellular transport, and cell division.

PUBLICATIONS

Hardy, J. H., et al. Multilevel summation method for electrostatic force evaluation, *J. Chem. Theory Comput.*, 11:766-779, 2015.

Perilla, J. R., et al., Molecular dynamics simulations of large macromolecular complexes, *Current Opinion in Structural Biology*, 31:64-74, 2015