THE COMPUTATIONAL MICROSCOPE

Allocation: NSF PRAC/19.6 Mnh
PI: Klaus Schulten[1]
Co-PIs: James C. Phillips[1] and John E. Stone[1]
Collaborators: Peijun Zhang[2], Angela Gronenborn[2], Tatyana Polenova[3], Christopher Aiken[4], Rebecca C. Craven[5], and Eva Nogales[6,7]

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

EXECUTIVE SUMMARY

In the petascale era, computational biology has redefined itself by connecting atomic-level descriptions of biological systems with cellular architecture and behavior. Molecular dynamics (MD) simulations serve as the computational microscope that empowers scientists with a tool to elucidate the dynamics and the underlying physical and chemical mechanisms of cellular processes. Due to its complementary role to experimental observation, MD unveils cellular organelles at atomic resolution.

INTRODUCTION

Recent developments in hybrid experimental methods, based on the revolutionary advance of electron microscopy, have led to previously unimaginable information on the 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 [1]. 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 is developed so that the biologically functional timescale in living cells can 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 (electron microscopy) 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

Solving the atomic-level structure of the mature HIV capsid [2] allowed us to study for the **first time** dynamic and structural properties of this multi-protein complex crucial for the biomedical community. Recently, experimental-computational studies demonstrated the influence of the human protein, Cyclophilin-A (CypA), in the dynamics of the capsid [3] and how hundreds of CypA bind to its surface [4]. Such properties may help scientists to understand better how the HIV capsid infects the host cells and could lead to new HIV therapies.

Virus infection and proliferation have been subjects of intense research for therapy development, targeting different maturation stages of the viruses. A strategy for preventing such proliferation, in particular, is to lock the viral particles in their immature, non-infectious state. Our investigation on the lattice model of the immature Rous sarcoma virus (RSV) [5], closely related to HIV, revealed a key source of dynamic and structural details presented on the recent experimental-computational characterization of the maturation process of the RSV virus [6].

The importance of simulating functional assemblies up to the level of complete capsids, opposed to isolated capsid proteins, was similarly demonstrated in a recent study. When bound to the HIV and hepatitis B virus (HBV) capsids, the drugs (PF74 on the HIV-1 and HAP1 on the HBV) caused changes not only in the vicinity of the binding sites but also on the global conformation of the capsids structure. The effect of PF74 was also evidenced on the dynamics of the processes taking place in distant regions of these structures [7].

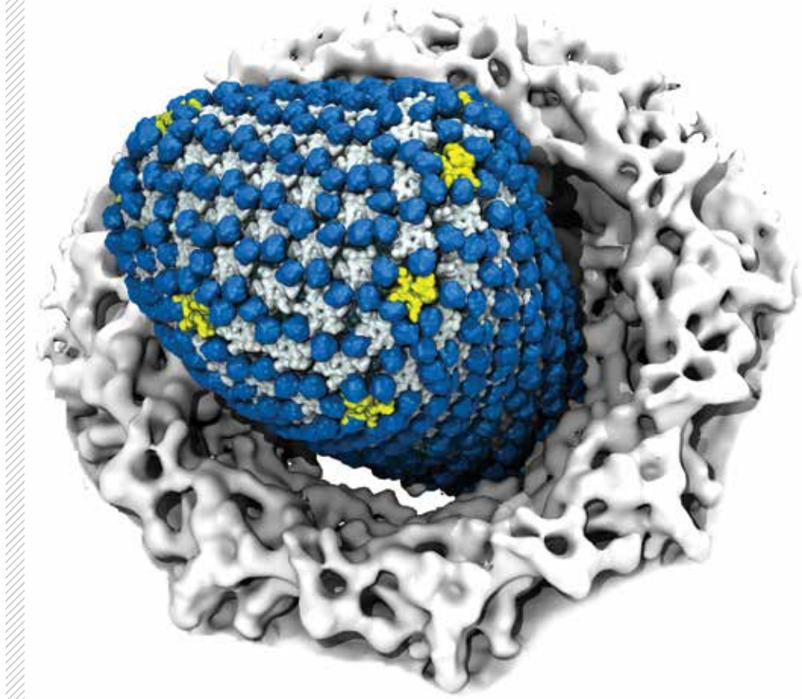

FIGURE 1: HIV-1 capsid docked to the human nuclear pore complex. HIV-1 exploits cellular pathways during infection, recruiting cyclophilin-A (blue) to go undetected in the cell.

Bacteria utilizes large, highly ordered clusters of sensory proteins, known as chemosensory arrays, to detect and respond to chemicals in their environment. Recently, we have integrated multi-scale structural data from experimental sources to computationally construct the **first atomic model** of the chemosensory array's molecular architecture [8]. Also, we identified a novel conformational change in a key signaling protein that is linked to chemotaxis function at the cellular level. 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 in different nucleotide binding states (crucial for switch between phases of assembly and disassembly). Our MD simulations suggested important, structurally dynamic events towards the microtubules' assembly and stability. Such simulations pave the way to understand the atomic details of the assembly and disassembly phases of the microtubules as well as the effect of anticancer drugs on microtubule dynamic instability.

WHY BLUE WATERS

Without Blue Waters and other petascale computing resources, projects involving large molecular systems like HIV, RSV, HBV, chemosensory array, and

microtubule would not be possible. These molecular systems are composed by several dozens of millions of atoms and must be simulated for long periods of time (microseconds). 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.

PUBLICATIONS AND DATA SETS

Goh, B. C., et al., Atomic modeling of an immature retroviral lattice using molecular dynamics and mutagenesis. *Structure*, 23 (2015), pp. 1414-1425.

Fraiberg, M., CheY's acetylation sites responsible for generating clockwise flagellar rotation in Escherichia coli. *Mol. Microbiol.*, 95:2 (2015), pp. 231-44.

Sisneros, R., C. Malone, A. Nonaka, and S. Woosley Chemical visualization of human pathogens: the Retroviral Capsids. *SC15 Exhibition, Huston, TX: High Performance Computing, Networking Storage and Analysis.*

Lu, M., et al., Dynamic allostery governs cyclophylin A-HIV capsid interplay. *PNAS*, 112: 47 (2015), pp. 14617-14622.

Cassidy, C. K., et al., CryoEM and computer simulations reveal a novel kinase conformational switch in bacterial chemotaxis signaling. eLife, (2015), 10.7554/*eLife*.08419.

Stone, J. E., Atomic detail visualization of photosynthetic membranes with GPU-accelerated ray tracing. *Parallel Comput.* 55 (2016), pp. 17-27.

Liu, C., et al., Cyclophilin A stabilizes the HIV-1 capsid through a novel non-canonical binding site. *Nat. Commun.*, 7 (2016), pp. 10714-10724.

Stone, J. E., Sherman, W. R. and Schulten K., Immersive molecular visualization with omnidirectional stereoscopic ray tracing and remote rendering. IEEE International Parallel and Distributed Processing Symposium Workshop (IPDPSW) (IEEE Computer Society, Chicago, Illinois, May 23-27, 2016).

Stone, J. E., et. al. , High performance molecular visualization: In-situ and parallel rendering with EGL. IEEE International Parallel and Distributed Processing Symposium Workshop (IPDPSW) (IEEE Computer Society, Chicago, Illinois, May 23-27, 2016).

Perilla, J.R., et. al., All-Atom Molecular Dynamics of Virus Capsids as Drug Targets. *J. Phys. Chem. Lett.* 7 (20126), pp. 1836–1844.

Heyrana, K., et al., Contributions of charged residues in structurally dynamic capsid surface loops to Rous sarcoma virus assembly. *J. Virol*, 90:12 (2016), pp. 5700-5714.

Stone, J. E., Early experiences porting the NAMD and VMD molecular simulation and analysis software to GPU-accelerated OpenPOWER platforms. International Workshop on OpenPOWER for HPC (ISC High Performance conference, Frankfurt, Germany, June 23, 2016).

THE RECYCLING MACHINERY OF THE CELL

Allocation: Illinois/500 Knh
PI: Klaus Schulten[1]
Co-PIs: Till Rudack[1] , Lela Vukovic[2]
Collaborator: Wolfgang Baumeister[3]

[1]University of Illinois at Urbana Champaign
[2]University of Texas at El Paso
[3]Max Plank Institute of Biochemistry

EXECUTIVE SUMMARY

While waste recycling became popular in our daily life more recently, living cells have mastered recycling of their protein content since their very beginning. Recycling of unneeded protein molecules in cells is performed by a molecular machine called 26S proteasome, which cuts these proteins into smaller pieces for reuse as building blocks for new proteins. Proteins that need to be recycled are labeled by tags made of poly-ubiquitin protein chains. The 26S proteasome machine recognizes and binds to these tags, pulls the tagged protein close, then unwinds it, and finally, cuts it into pieces.

Despite its substantial role in the cell's life cycle, the proteasome's atomic structure and function remain elusive. Employing a combination of computational techniques implemented using nanoscale molecular dynamics (NAMD) along with cryo-electron microscopy (EM) data, we obtained an atomic structure of the human 26S proteasome and investigated the mechanism underlying substrate recruitment and unfolding.

FIGURE 1: The recycling system of the cell. The 26S proteasome is the key player of the human protein recycling system. The first structure of the human 26S proteasome obtained through integrative modeling utilizing Blue Waters will lead to breakthroughs in understanding its detailed function and will play a pivotal role in the development of the 26S proteasome as a drug target for molecular disease therapies.

INTRODUCTION

Recycling of proteins by degradation is vital for a variety of essential cellular processes, including protein quality control, cell cycle regulation, adaptive immune response, and apoptosis. The 26S proteasome is responsible for the vast majority of regulated intracellular protein degradation and is an important drug target for multiple diseases, including cancer, neurodegenerative diseases, and immuno-inflammatory disorders. The 26S proteasome is an adenosine triphosphate (ATP) hydrolysis driven 2.5 MDa molecular machine that recruits, unfolds, and degrades poly-ubiquitin tagged proteins through a complex interaction clockwork of over 60 known protein subunits (Fig. 1).

Despite its substantial role in the cell's life cycle, the proteasome's atomic structure and function remained elusive. However, recent developments in hybrid experimental methods based on the revolutionary advance of electron microscopy, together with improvements in real space refinement methods [1], have led to previously unimaginable