Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Molecular dynamics simulations of large macromolecular complexes

Juan R Perilla[1,2], Boon Chong Goh[1,2], C Keith Cassidy[1,2], Bo Liu[1,5], Rafael C Bernardi[1,3], Till Rudack[1], Hang Yu[1,5], Zhe Wu[1,4] and Klaus Schulten[1,2,3,4,5]

Connecting dynamics to structural data from diverse experimental sources, molecular dynamics simulations permit the exploration of biological phenomena in unparalleled detail. Advances in simulations are moving the atomic resolution descriptions of biological systems into the million-to-billion atom regime, in which numerous cell functions reside. In this opinion, we review the progress, driven by large-scale molecular dynamics simulations, in the study of viruses, ribosomes, bioenergetic systems, and other diverse applications. These examples highlight the utility of molecular dynamics simulations in the critical task of relating atomic detail to the function of supramolecular complexes, a task that cannot be achieved by smaller-scale simulations or existing experimental approaches alone.

Addresses
[1] Beckman Institute, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[2] Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[3] Energy Biosciences Institute, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[4] Center for the Physics of Living Cells, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[5] Center for Biophysics and Computational Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

Corresponding author: Schulten, Klaus (kschulte@ks.uiuc.edu)

Current Opinion in Structural Biology 2015, **31**:64–74

This review comes from a themed issue on **Theory and simulation**

Edited by **Claire Lesieur** and **Klaus Schulten**

For a complete overview see the Issue and the Editorial

Available online 4th April 2015

http://dx.doi.org/10.1016/j.sbi.2015.03.007

Introduction

The essential conundrum of modern biology, namely the question of how life emerges from myriad molecules whose behavior is governed by physical law alone, is embodied within a single cell — the quantum of life. As illustrated in Figure 1, the rise of scientific supercomputing has allowed for the study of the living cell in unparalleled detail, from the scale of the atom [1••,2] to a whole organism [3–5] and at all levels in between [6]. In particular, the past three decades have witnessed the evolution of molecular dynamics (MD) simulations as a 'computational microscope' [7], which has provided a unique framework for the study of the phenomena of cell biology in atomic (or near-atomic) detail.

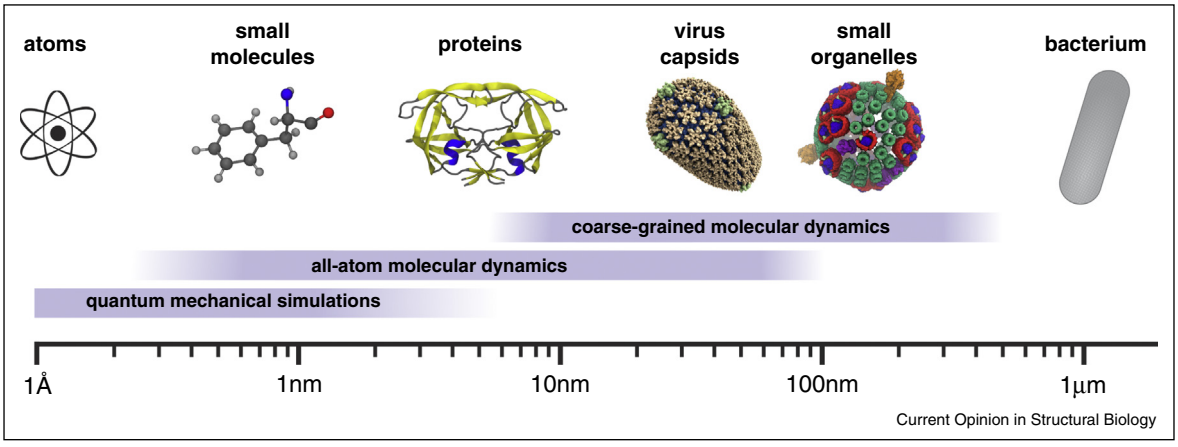
Now, in the era of petascale computing, high-performance MD software packages such as NAMD [8], GROMACS [9], and LAMMPS [10] are being optimized for scaling to an ever-increasing number of cores on cutting-edge computing hardware [2,11,10], enabling the investigation of previously unfathomable biological phenomena through the use of large-scale atomistic simulations. Moreover, the development of computational tools such as molecular dynamics flexible fitting (MDFF) [12,13] are forging an intimate connection between experiment and theory, informing the construction of atomic-level models of large-scale, supramolecular complexes through a synthesis of multiscale experimental data from cryo-EM, NMR spectroscopy, and X-ray crystallography. Complementary to all-atom MD simulations, the development of force fields for coarse-grained MD (CGMD) simulations continues to be a popular source of techniques which favor computational efficiency over atomic and chemical accuracy, permitting simulations on even larger time and length scales [14].

This opinion will focus on the ways in which large-scale MD simulations are having a profound impact in numerous diverse scientific endeavors. From the treatment of disease and development of drugs [15•,16•] to the fabrication of novel biomaterials [17•] and creation of bio-based renewable energy sources [18], large-scale MD simulations are helping to achieve a fundamental understanding of living organisms. Taken together, the work reviewed here demonstrates the maturity of the MD apparatus as a tool to progress basic science and the investigation of the molecular makeup of life.

Large-scale MD simulations of viruses

Viruses are parasitic life-forms that replicate by hijacking resources present in the cells they infect. Because of their small size compared to cells (20–1500 nm scale), observation of the viral particle during different stages of the

Figure 1



Characteristic length-scales currently associated with varying levels of description in biomolecular simulations. *Ab initio* and semi-empirical quantum mechanical calculations permit the study of chemical reactions in electronic detail within single molecules and small proteins while all-atom and coarsed-grained molecular dynamics simulations allow for the study of biological phenomena from the individual protein level to large subcellular organelles, and at all levels in between.

replication cycle is mostly limited to electron microscopy. Yet, virus particles are large in size for all-atom simulations (see Figure 2) and as a result most studies at the atomic level have been limited to isolated virus proteins or subfragments of a viral particle or capsid [19]. The satellite tobacco mosaic virus (STMV) became the first complete virus to be investigated through all-atom MD simulations [20]. Since then, MD programs have become capable of simulating systems of even larger size and complexity [1••,11,21•], thus allowing the study of viral particles up to two orders of magnitude larger in atom count than STMV [1••].

High resolution structures of symmetrical virus capsids like poliovirus, southern bean mosaic virus and satellite tobacco necrosis virus have been available for several years, leading to routine investigations using MD simulations [22,24–26]. More recently, MDFF has been applied to elucidate the structures of yet larger and more complex virus capsids in their native environments [27,15•,1••,28]. For instance, MDFF was instrumental in the structural determination of the HIV-1 core [1••], a polymorphic capsid with no apparent symmetry. The 64-million-atom MD simulation of the mature HIV-1 capsid should enable the characterization of complex interactions between host cell factors and the assembled capsid lattice [29], thus providing an unexploited framework for the development of novel drugs targeting HIV-1. Similarly, MDFF's application to the Rabbit hemorrhagic disease virus (RHDV), a 10-million-atom MD calculation, improved the overall fitting of the crystal structure to the cryo-EM density, leading to the development of a vaccine [15•].

Spontaneous assembly of the immature and mature forms of the capsid from viral nucleic acids and proteins is an essential step for the replication cycle of a virus. However, the rate at which assembly occurs prevents a molecular-level description of the process by experimental means, making it an attractive task for computational modeling. The virus assembly kinetics and pathways can be computationally determined using stochastic kinetics [30], elastic networks [31], simulation-based data fitting [32], and CGMD simulations [23•33,30,34]. In fact, by means of CGMD simulations, optimal configuration of the viral genome has been shown to be essential for proper assembly of the capsid [23•,33]. In the particular case of simian virus 40 (SV40), two pathways have been observed during assembly, with each of the pathways sampling different intermediate states depending on the strength of both ionic solution and protein–protein interactions [23•]. Coarse-grained simulations have also been performed to investigate the self-assembly process of the HIV-1 mature capsid [35,36,1••], allowing the identification of a trimer-of-dimers structure as a fundamental step during assembly of hexagonal lattices in a crowded environment [35,36]. However, due to the lack of structural information available at the time for informing the CG models, both studies neglected explicit interactions at the trimeric interface of the lattice, an interface now known to play an essential role in capsid curvature, assembly, and stability as well as infectivity *in vivo* [1••].

In order to become infectious, entire viral particles undergo a global structural rearrangement known as maturation. Such a maturation process is thought to follow two different pathways: disassembly–reassembly