ATOMISTIC SIMULATION OF A PROTOCELL

Allocation: Illinois/994 Knh
PI: Emad Tajkhorshid[1]

[1]University of Illinois at Urbana–Champaign

EXECUTIVE SUMMARY

Computational biology has been critical in achieving a molecular understanding of the dynamics and function of proteins, nucleic acids, and biological membranes. Most research in this field has been limited to single-protein characterization owing to limits in computational power and analytical methods. However, to truly understand the mechanics of biological systems, proteins need to be studied in the broader context of a cell. This project aims to spearhead a new generation of computational biology research that seeks insight from molecular dynamics (MD) simulations to study cellular-scale processes in atomic detail. The immensely complex nature of a cell belies the elementary physics that govern it, and by using MD simulations a rigorous connection can be made between the underlying physical laws and the emergent phenomena they generate. By pioneering new, large molecular simulations, researchers can begin to gain insight into and to quantify mesoscale biological phenomena.

RESEARCH CHALLENGE

The living cell represents a spatially complex and highly regulated arrangement of molecules whose coordinated motions and activities underlie all the processes that allow it to grow, reproduce, and carry out a wide spectrum of cellular functions. Recent advances in experimental techniques have targeted the identification of the positions of subcellular organelles, ribosomes, and macromolecules such as proteins, messenger RNA, and DNA at high resolution [1]. This has in turn fed a growing demand for next-generation computational tools that allow researchers to construct realistic, cellular-scale structural models at a range of resolutions; to resolve molecular functional states; and to simulate their stochastic, time-dependent behavior.

Biological membranes constitute the most fundamental component of cellular and subcellular structures, defining the active boundaries not only between the cell and its environment but also between different compartments within the cell. Diverse in form, structure, and lipid composition, they are highly heterogeneous environments containing lipids and proteins that act in concert to regulate the flow of information and materials into and out of the cell [2,3]. Modeling membranes at cellular scales is an emerging challenge that remains to be addressed and, therefore, is one of the critical steps in modeling the cell. The research team has a long tradition in modeling and simulating membranes and membrane proteins, particularly emphasizing the importance of including the natural environment in the simulations [4–8].

METHODS & CODES

The MD simulations were performed with NAMD [9], a highly parallelized, GPU-accelerated, publicly available MD program with demonstrated scalability to hundreds of thousands of processors for both single- and multiple-replica simulations. All-atom MD simulations rely on the accurate integration of the equations of motion for all atoms of the system. The total potential energy of the system was described by the CHARMM36m force field [10,11]. Periodic boundary conditions were used to avoid surface effects at the simulated system's boundary, allowing the efficient computation of nontruncated electrostatic interactions by the fast Fourier transform-based particle-mesh Ewald method [12].

RESULTS & IMPACT

The research team designed, developed, and tested a protocol for the construction of spherical cellular envelopes of any lipid/protein composition, which can be divided into three stages. The first step in the designed workflow was to use experimentally derived or user-specified cell diameters to directly determine the surface area of the inner and outer leaflet of the cell envelope and to accurately calculate the number of lipids needed to construct both leaflets. The membrane shell was then generated using the Fibonacci sphere algorithm [13] to approximately solve the Tammes problem [14,15] for the thousands of lipids required to create a micrometer-scaled membranous protocell. Next, each protein was assigned a unique position on the spherical surface. In order to do this, coarse-bead approximations of each protein were generated using a hierarchical clustering algorithm. These were then simulated using a force field that properly oriented and localized them to the intended membrane surface. Lastly, the membrane and the proteins were merged, which involved removal of requisite lipids from both the inner- and outer-membrane leaflets to ensure conservation of surface area and avoid clashes between lipid and protein molecules.

The number of lipids that each protein displaced was carefully derived from conventional (smaller) protein-embedded lipid bilayer simulations and dictated how many lipids were removed from the spherical membrane. Lipids were hierarchically eliminated based on their degree of impingement into the intended protein volume and any remaining potential protein–lipid clashes were resolved via a brief grid-steered simulation of the membrane that forces any offending lipids from the protein space. This procedure ensured that the exact number of lipids matching the cross-sectional area of the protein (calculated independently for the inner and outer leaflets) were removed. Finally, the membrane and proteins were merged together and solvated, and the final system was used for MD simulations.

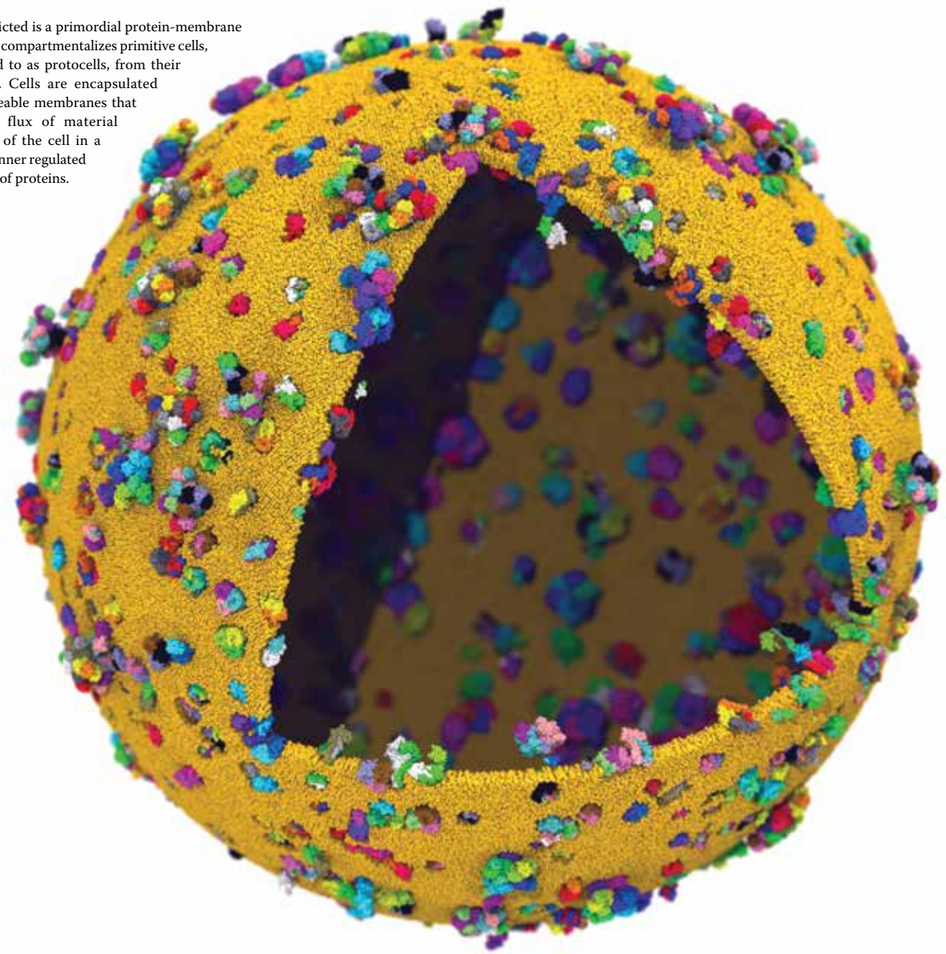

Figure 1: Depicted is a primordial protein-membrane structure that compartmentalizes primitive cells, often referred to as protocells, from their environment. Cells are encapsulated by semipermeable membranes that regulate the flux of material into and out of the cell in a concerted manner regulated by a diversity of proteins.

The simulations performed on Blue Waters for this project informed areas of refinement for the research team's assembly protocol, particularly for the removal of lipids during the incorporation of proteins and solvation procedures. The final simulations indicated the constructed protocell systems are stable and that the assembly protocols are adequate. The project presents the most carefully crafted and stable protocell at full atomic detail (Fig. 1).

WHY BLUE WATERS

For this project, the researchers performed MD simulations with NAMD of systems measuring one billion atoms in size, which can only be achieved on a petascale computing platform such as Blue Waters. The GPU accelerators on Blue Waters confer a significant boost to simulation speed (two to three times) versus the non-accelerated CPU-only nodes. NAMD has been extensively tested and optimized for Blue Waters, making use of the high-speed node interconnect and showing sustained petascale performance across hundreds, even thousands, of nodes. The team's benchmarks demonstrated an efficient scaling performance (greater than 80%) while using up to 2,048 Blue Waters GPU (XK7) nodes for a one-billion-atom system.