We investigated the detailed molecular mechanism of activation of TRPV1. We found that dynamically controlled solvation of hydrophobic protein pockets is the key determinant of activation. We are now in the process of extending this investigation to address the relevant issue of how this structural transition is triggered by environmental stimuli such as temperature, osmotic pressure and binding of PIP_2 lipids.

METHODS & RESULTS

We performed extensive multi-microsecond molecular dynamics (MD) simulations on several TRPV1 systems, varying the initial setup and the equilibration protocol for the purpose of characterizing stable conformational states close to the experimental structures.

Binding mode of agonists

Initially, we focused on the binding mode of agonists, an issue that was not solved by the cryoEM investigation. We used the information contained in the experimental electron density maps to determine the binding mode of capsaicin and resiniferatoxin. In collaboration with the lab of Tibor Rohacs at Rutgers New Jersey Medical School, we validated the theoretical predictions via mutagenesis experimentally. By these results, we characterized the structural rearrangements entailed by binding of each ligand to the pocket (Fig. 1).

Open to close transition

We then started MD simulations of the capsaicin-bound state in a standard POPC (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine) bilayer to explore the mechanism of channel opening. At the end of the 750 ns MD trajectory, we analyzed the root-mean-square deviation (RMSD) from the initial capsaicin-bound structure and from the open structure. Surprisingly, we found that two subunits out of four diverged from the initial state and relaxed to the open state. Consistently, after approximately 400 ns of simulation, we observed a sudden increase in the pore radius and a simultaneous hydration of the so-called hydrophobic gate. We thus compared the radius and the hydration profile of the resulting open state to those of the closed state and found significant differences (Fig. 2). By performing metadynamics simulations, we finally ascertained that the conformational state obtained from the simulation is conductive for sodium (Fig. 2C).

WHY BLUE WATERS

The project relied crucially on a quantitative description of complex processes occurring on time-scales of several microseconds in large membrane-protein assemblies with a typical size of approximately 300,000 atoms. The system is constituted by an ion channel, a model lipid bilayer, and an electrolyte solution and was simulated under different conditions, i.e. in presence or in absence of a ligand. The capabilities of Blue Waters turned out to be key to the success of this computationally intensive project.

NEXT GENERATION WORK

The complexity behavior of ion channels relies in part on the peculiar environment of the lipid membranes present in neurons. The insight developed during this project will enable computational modeling of the response of a channel in a neuron-like environment; ultimately, the goal is to develop a computational framework to quantitatively address the following question: how does the nervous system transduce and transmit the information about noxious environmental stimuli?

PUBLICATIONS AND DATA SETS

Elokely, K., et al., Understanding TRPV1 activation by ligands: Insights from the binding modes of capsaicin and resiniferatoxin, *PNAS,* 113:2 (2016), pp. E137-E145, doi:10.1073/pnas.1517288113

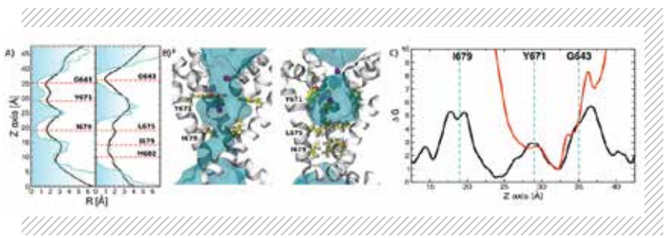

FIGURE 2: Open and closed states of TRPV1. A) The radius of the pore (black line) is shown along with the water density profile (blue shading) for the open (left) and closed (right) conformation. B) Three-dimensional shape of the pore (blue shading) is shown together with pore-lining residues. C) Free-energy profile for the permeation of a sodium ion as calculated form metadynamics for the open (black) and closed (red) confomations. Figure adapted from ref. (Kasimova, 2015).

RIBOSOME BIOGENESIS IN REPLICATING CELLS

Allocation: Illinois/606 Knh
PI: Zaida Luthey-Schulten[1]
Co-PIs: Tyler Earnest[1] and John Cole[1]
Collaborator: Thomas Kuhlman[1]

[1]University of Illinois at Urbana-Champaign

EXECUTIVE SUMMARY

Ribosomes—the cellular machines responsible for making proteins—are complexes of nucleic acids and proteins. The ways in which the living cell manages their production and assembly is of deep biological importance and is highly conserved among all bacteria. Here we extend a recent spatially resolved whole-cell model of ribosome biogenesis in cells of a fixed volume [1] to include the effects of DNA replication, growth, and cell division [2]. Cell cycle parameters describing the replication schedule were obtained by analyzing single cells from a series of *Escherichia coli* strains with fluorescently labeled genes distributed evenly throughout the chromosome. For cells with a two hour generation time, replication was initiated 42 minutes into the cell cycle and completed after an additional 42 minutes. All of the biological processes of replication, transcription, translation, and ribosome assembly are described regarding reaction-diffusion master equations and solved stochastically using Lattice Microbes v2.3a software package [3-5].

INTRODUCTION

In bacteria, ribosomes account for approximately one-fourth of the cellular dry mass and the majority of the total RNA [1]. The ribosomes' role in protein synthesis couples them to essentially every process within the cell. It can be tempting to think of the bacterial cell as a finely tuned machine for building ribosomes. Previously, we published a model of ribosome assembly in the cell based on kinetic and thermodynamic data [2-3]. Although unprecedentedly complete, the model did not account for some of the most basic functions of the cell—DNA replication, growth, and cell division. Duplication of the chromosome affects the copy number distribution of gene products since the rate of transcript production effectively doubles [4]. However, to correctly capture the shape of the distribution the behavior over the interval where the mRNA copy number relaxes to the new steady state under the doubled gene dose must be accounted for [8-9]. The changing cell volume also impacts the copy number dynamics due to the dependence of reaction rates on volume. To account for this, we have augmented this model by deterministically modeling cell growth and DNA replication using experimental data from our collaborator, Thomas Kuhlman. *E. coli* follows a simple three-period cell cycle based on the timing of DNA replication and cell division (Fig. 1). Using estimates of the durations of these periods as well as the mean cell length at division computed from experiments performed by our collaborator, we extended our model to include gene replication and cell growth through deterministically modifying the cell geometry, operon locations, and copy numbers over the course of the cell cycle.

METHODS & RESULTS

Previously, we constructed a kinetic model of the biogenesis of the ribosomal small subunit (SSU), including the transcription of rRNA and mRNA coding for the ribosomal proteins and transcription and translation of mRNA [1]. This model was embedded in a spatial model of *E. coli*, derived from cryo-electron tomograms [10], which takes into account the geometry of the cell and its compartmentalization into four distinct regions: extracellular, membrane, cytoplasm and nucleoid (Fig. 1). This system was simulated on Blue Waters using the Lattice Microbes v2.3a software package, which includes an improved algorithm for computing reaction propensities [6] (Fig. 3).

Our simulated cell was allowed to grow from 2.4 to 4.8 μm over its 120-minute cell cycle. Genes are replicated in the simulation at times determined from the cell cycle parameters and their position in

the genome. Compared to a fixed-volume version of this model (without growth and cell division), the initial and final species counts are practically identical for all protein types, ribosomal subunits, and translating ribosomes. However, since the volume expansion of the cell proceeds at the same rate as the new species are created, the cell can maintain a constant concentration of ribosomes and ribosomal protein. Finally, the importance of a stochastic, spatially resolved representation is emphasized yet again due to the limited copy number and lifetimes of the assembly intermediates.

WHY BLUE WATERS

Our simulations are computationally intensive and require high-performance graphics processing unit (GPU) accelerators. The stochastic nature of these simulations requires multiple realizations of the cell cycle to properly investigate the statistical variation of the cell's phenotype. These simulations are I/O intensive: on average, 20 GB of data is generated for a single realization of a cell cycle. This demands both high-capacity storage as well as high bandwidth I/O. Blue Waters provided the technical resources necessary to study this system. The recent introduction of a Blue Waters Python environment (BWPY) by the staff allowed us to streamline our development/simulation cycle by providing an up-to-date and complete Scientific Python stack optimized for Blue Waters. Since our code relies on modern Python features, this relieved us from the task of building our Python environment optimized for a supercomputing environment.

NEXT GENERATION WORK

In the next generation Track-1 systems, we hope to build a highly realistic *in silico* bacterium that will incorporate spatially resolved signaling, transcription, replication, translation and metabolic networks which include chemical species at densities spanning many orders of magnitude. The computational challenge of simulating the numerous metabolites and reactions needed to describe these processes will require the use of multiple nodes as well as multiple GPU accelerators per node for a realization of a single cell. Next-generation GPUs allow significantly faster time-to-solution (Fig. 2) and will allow the study of longer timescales while collecting the requisite sample size for proper statistical analysis. This challenge will necessitate the continual development of our Lattice Microbes Message Passing Interface code and implementation of hybrid deterministic–stochastic treatments of species allowing for a density scale separation, which will avoid the use of computationally intensive stochastic treatments unnecessary for highly abundant species.

PUBLICATIONS AND DATA SETS

Earnest, T.M., et al., Toward a Whole-Cell Model of Ribosome Biogenesis: Kinetic Modeling of SSU Assembly, *Biophys. J.*, 109:6 (2015), pp. 1117–1135. doi:10.1016/j.bpj.2015.07.030

Earnest, T.M., et al., Ribosome biogenesis in replicating cells: Integration of experiment and theory. *Biopolymers*, (2016). doi:10.1002/bip.22892

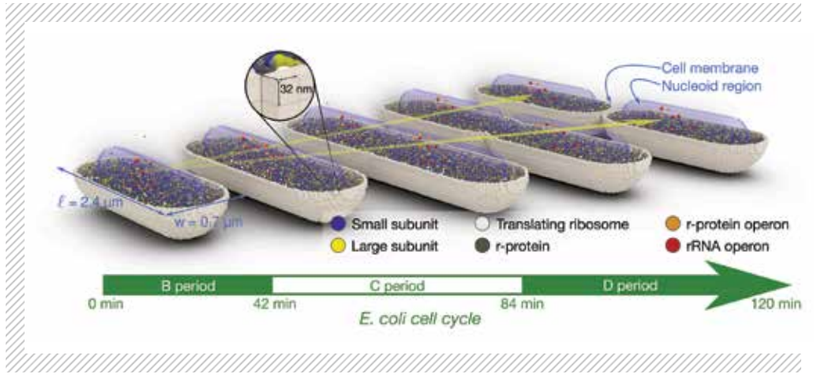

FIGURE 1: Schematic of whole-cell simulations of ribosome biogenesis in slow growing *E. coli*. The simulation is performed on a 32 nm lattice of dimensions 32×32×192, with four compartments representing the extracellular, cell membrane, cytoplasm, and nucleoid regions. The cell cycle is divided into three periods: the time after cell division but before the initiation of DNA replication (B period), the period of DNA replication (C period), and the lag time between the termination of replication and completion of cell division (D period). Using experimental estimations of the duration of these periods, we are able to schedule the replication of the ribosomal operons (red and orange spheres) during the cell cycle. These operons are placed along the long axis of the cell based on their location in the genome. The cell grows from 2.4 µm to 4.8 µm over 120 minutes, increasing its particle count from ~50,000 to ~100,000, while new ribosomal operons are added at times derived from the cell cycle parameters and the genomic distance from the origin of replication, which are moved along the long axis of the cell such that the operons are found in the same relative positions in the daughter cells as in the mother cell (yellow arrow).

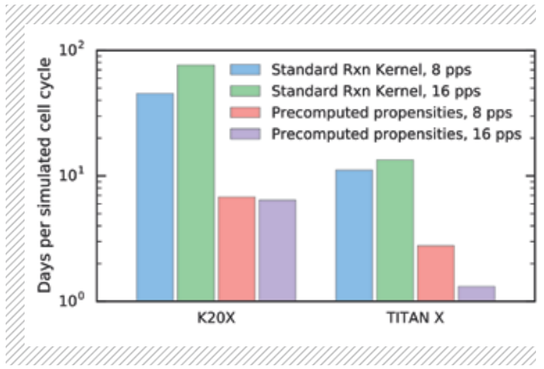

FIGURE 2: Compute time necessary to simulate a 120-minute cell cycle. The standard reaction kernel computes the reaction propensity for each possible reaction, whereas the precomputed propensities kernel uses precomputed values and loops over the number of particles per site. The latter method is significantly faster due to the unusually large number of reactions (800) in the model. By increasing the total number of particles from 8 to 16 per lattice site, a speed-up is achieved in spite of the increased time necessary to transfer the lattice data. Since this model can attain high particle densities, the greater lattice capacity reduces the rate of overflows, which occur when a site has exceeded its capacity and the new particle must be placed in a nearby site. An approximate 6× speed-up switching from the Kepler architecture (K20X, 6 days) to Maxwell (TITAN X, 1 day) was observed.