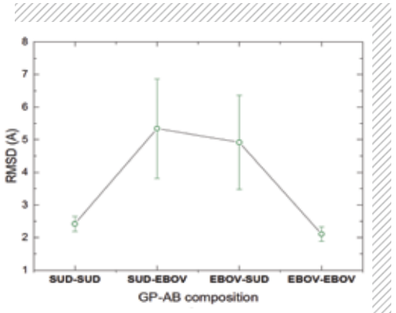
FIGURE 2: RMSD (magnitude of structural fluctuations, a measure of structural instability) derived from MD simulations when Sudan strain antibody is matched with Sudan virus, when Sudan antibody is matched with Zaire virus, when Zaire strain antibody is matched with Zaire virus, and when Zaire antibody is matched with Sudan virus. This image shows proof-of-concept that molecular dynamics is sufficiently sensitive to distinguish between effective binding (that can overcome the infection by preventing viral entry into cells) and ineffective binding.




antibody complexes [6]. However, this description of the glycoprotein-antibody competition does not in itself lead to a predictive model for how the virus will evolve and what change in the antibody will be effective against the evolved viral protein.

## METHODS & RESULTS

We began by considering how to construct a predictive model for how the virus is likely to evolve, and what alterations in the sequence of a binding region of the antibody would most effectively counter the viral mutation(s) and restore the ability of the antibody to bind the glycoprotein. To predict likely mutations, we used existing statistical data on the likelihood of particular substitutions, as embodied in a "substitution matrix" in which each element corresponds to a relative probability of an amino acid substitution [7]. To predict effective responses to viral mutations, we used existing statistical data on amino acids that interact favorably at protein-protein interfaces [8]. Finally, we used MD simulations of the mutated glycoprotein-antibody complex to test the statistical prediction by computing the effects of the postulated mutations [9].

The starting points for the simulations were structures of Ebola glycoprotein complexed with antibody fragments that were known to prevent infection successfully. We then mutated the glycoprotein, which invariably resulted in degradation of binding energy between glycoprotein and antibody. Following that, we used the databases of favorable amino acid interactions to make educated guesses as to mutations on the antibody. In the majority of cases, we were able to re-engineer the antibody to bind the viral protein as well or better than the wild type. In those cases where we did not succeed in doing that, we believe that we would have succeeded if provided with more computertime.

## WHY BLUE WATERS

We could not have done the project without the sheer computational power of Blue Waters.

## NEXT GENERATION WORK

Our long-term goal is to establish our approach as a standard for the design and redesign of synthetic antibodies against viral infections so that humankind can prevail in the evolutionary arms race against evolving viral pathogens.

## PUBLICATIONS AND DATA SETS

Farimani, A.B., et al., Computational approach to designing antibody for Ebola virus. *Biophyss J.*, 110:3 (2016)537a.

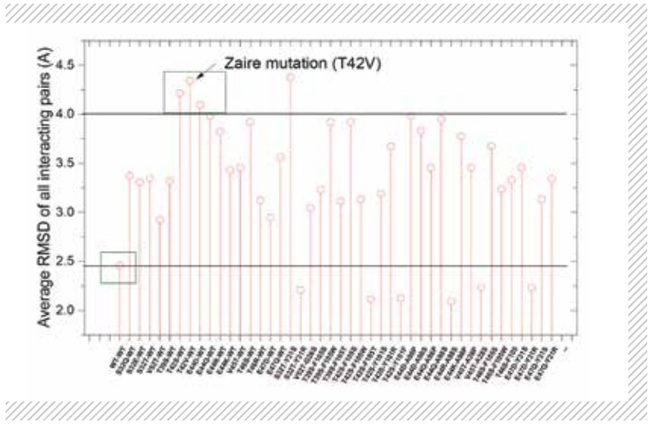



FIGURE 3: RMSD for selected mutants of the Sudan virus and mutated antibodies. The symbols that fall below 2.5 angstroms are actually binding better than wild type. Antibody redesign succeeded by that criterion for approximately half of the mutant glycoproteins.





# UNVEILING ALLOSTERIC PATHWAYS IN ION CHANNELS

**Allocation:** NSF PRAC/5.60 Mnh
**PI:** Michael L Klein[1]
**Co-PI:** Vincenzo Carnevale[1]

[1]Temple University

## EXECUTIVE SUMMARY

Transient receptor potential (TRP) channels are central to environmental sensation in animals, fungi, and unicellular eukaryotes. All known TRP channels are nonselective cation channels that open in response to a wide array of factors. Clarifying how TRP channels convert physical and chemical stimuli from the environment into the allosteric signals underlying channel activation is key to understanding how they control cell excitability in both physiological and pathological conditions. Their relevance in the molecular pathways that mediate pain makes them promising targets for novel classes of analgesics (medicines that relieve pain). Building on the structural information made recently available for transient receptor potential cation channel subfamily V member 1 (TRPV1), thanks to a series of cryo-electron microscopy (CryoEM) experiments, we performed free energy (metadynamics) simulations on models of TRPV1 embedded in a lipid bilayer. Harnessing the computation capabilities of Blue Waters, we explored several pathways of activation and characterized ion channel conductance and selectivity. Our calculations reveal a **novel** mechanism for sensing temperature and osmolality.

## INTRODUCTION

A fit cell must perceive and comprehend the conditions of its inner and outer worlds, integrating diverse and transitory physicochemical stimuli into concerted cellular decisions. For this reason, the membranes of even the simplest bacteria are studded with ion channel proteins that detect cellular conditions and translate them into electrochemical information via gated ionic conduction [1].

In eukaryotes, the complexity of cellular life has taken this requirement to its apex. Accordingly, natural selection has elaborated on the ion channel, producing an impressive array of polymodal cellular sensors, the TRP channels [2]. All TRP channels detect multiple physicochemical stimuli, with some overlap among the eight extant TRP subfamilies. However, the response to each stimulus varies substantially from channel to channel, presumably dictated by heterogeneous and subfamily-specific intra- and extracellular domains [3]. Indeed, since their divergence from the voltage-gated potassium (Kv) channel superfamily over a billion years ago, TRP channel proteins have maintained a tetrameric six-transmembrane (6-TM) architecture and little else. The (TRPV1) or vanilloid receptor 1 is a polymodal mammalian nociceptive integrator [4] abundantly expressed in the free nerve endings of primary pain-sensing afferent Aδ and C fibers [5]. Structurally, the TRPV1 channel is a homotetramer, symmetrically organized around a solvent exposed central pore. Each subunit is formed by six transmembrane helices (S1–S6) with the channels' N- and C-termini located in the intracellular medium [6].

TRPV1 is activated by a wide range of proinflammatory and proalgesic mediators [7]; including temperatures above 43°C, external pH, bradykinin, anandamide, arachidonic acid metabolites, jellyfish and spider toxins, vanilloid and others. The scope of the TRPV1 pharmacological spectrum [8-10] is mainly in the area of analgesics: novel painkillers could be either TRPV1 agonists or antagonists. Moving forward toward the rational drug design of TRPV1 modulators requires a basic understanding of how known ligands trigger the closed to open transition in TRPV1.

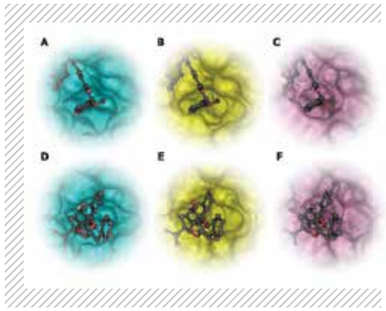



FIGURE 1: Optimal docking poses of capsaicin and resiniferatoxin in the vanilloid binding site of TRPV1. Shown are: A) capsaicin in the apo structure; B) capsaicin in the TRPV1-capsaicin complex; C) capsaicin in TRPV1-resiniferatoxin complex; D) resiniferatoxin in the apo structure; E) resiniferatoxin in the TRPV1-capsaicin complex; F) resiniferatoxin in the TRPV1-resiniferatoxin complex. Adapted from ref. (Elokely, 2015)

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

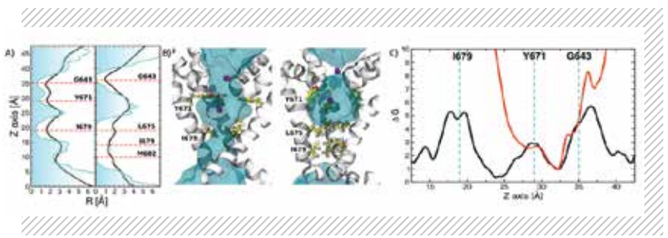



FIGURE 2: Open and closed states of TRPV1. A) The radius of the pore (black line) is shown along with the water density profile (blue shading) for the open (left) and closed (right) conformation. B) Three-dimensional shape of the pore (blue shading) is shown together with pore-lining residues. C) Free-energy 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