

## Research Statement

My laboratory is keenly interested in discovering the rules of chemistry shaping cellular function. We fulfill this passion by developing computational tools founded on fundamental chemical physics to probe cell function from single atoms to entire tissues (see Fig. 5). We use molecular simulations, statistical physics, computer vision and numerical algorithms to understand biochemical phenomena driving cardiac and neurological function. Our predominant focus is to create practical computational tools to improve researchers' ability to probe how macrophages, immune cells that mediate innate immunity, perform their pro- and anti-inflammatory roles *in vivo*. Our progress has already resulted in my lab securing almost \$2M in university, private and federal grants as well as over a dozen publications toward addressing the following challenges:

- How to engineer calcium-binding proteins to blunt gene activation associated with chronic inflammation
- Determine how signaling molecules' interactions with their surroundings impact macrophage activation
- Predict how macrophage intracellular signaling networks are triggered by extracellular ligands *in vivo*

In answering these questions, I am pursuing a long term goal of using mathematics, physical chemistry and computation to understand how biological signaling pathways are controlled at molecular through cellular levels, how they adapt to stress and dysfunction and how they can be modulated to treat human disease.

### Theme 1: Engineering proteins to selectively control calcium signaling processes.

Calcium is an intracellular signaling molecule that proteins used to communicate, convert chemical energy into mechanical work and shape macrophage phenotypes. It is unsurprising therefore that calcium dysregulation in macrophages correlates with many diseases including Alzheimer's and chronic pain. This fact has motivated considerable interest in developing therapeutics that target dysregulated calcium signaling. Engineering calcium binding proteins is one such strategy that has emerged in recent years and has been an active area of research in my lab. However, despite decades of theoretical descriptions of calcium-binding and the availability of hundreds of high-resolution calcium binding protein structures, engineering these proteins to tune their properties has largely been a guess-and-check process with limited successes.

We **hypothesize** that calcium affinity, if not binding kinetics, can be modulated by orders of magnitude through controlling proteins' changes in structure or solvent-exposed surface that occur during calcium binding. **We are approaching this challenge through an innovative method that combines statistical physics models of ion binding and protein/protein interactions with 3-D protein structures, binding poses and interaction energies obtained from atomistic resolution molecular simulations (Fig. 5a).**<sup>1-3</sup> The statistical physics models are portrayed in Fig. 1 and include a Mean Sphere Approximation theory that relates the packing of hard, charged atoms in binding sites to ion affinity (top) as well as worm-like chain that estimates a flexible peptides' radial distribution to its sequence length (bottom). This hybrid approach allows us to obtain accurate chemical potentials of binding and ligand distributions which are relevant to protein function. We used this approach to predict binding affinities in  $\beta$ -parvalbumin, a candidate for cardiac gene therapy, altered selectivity in variants of the calcium pump SERCA, and the regulation of calcineurin by its C-terminal domain.<sup>1-4</sup> Ultimately, these studies suggest that combining statistical physics with molecular structural information could help screen candidate mutations for engineering augmented or attenuated calcium-dependent functions of proteins like  $\beta$ -parvalbumin.

One of our most exciting recent developments has been resolving aspects of a poorly understood mechanism in the calmodulin (CaM)-dependent regulation of calcineurin. Calcineurin is involved in priming gene transcription in activated macrophages. This often happens in response to fluctuations in intracellular calcium content that arise from activating receptors on the macrophage plasma membrane. This process is important in neurological development, but can contribute to chronic inflammation under pathological conditions. Given the cytotoxicity of most calcineurin inhibitors that are candidates for blunting dysfunctional gene activation, there is an interest in identifying other target-able features, such as the mechanism by which calcineurin regulates itself through binding its intrinsic regulatory domain. However, the structure of regulatory domain has not been determined, which undermines structure-based inquiry. Intriguingly, there is structural evidence that the calcineurin regulatory domain folds upon calmodulin (CaM) binding, and further, that there is a region distal CaM binding region (a 'distal helix') that must adopt a helical configuration to promote enzyme activity. We **hypothesized** based on our progress with modeling short fragments of the calcineurin regulatory domain<sup>5</sup> that complete helical fragments could bind CaM and thereby pull the regulatory domain away from the enzyme's catalytic site. We devised an innovative hybrid approach combining molecular simulations with a worm-like chain model (from polymer physics) of its regulatory domain to predict CaM-compatible binding poses and their impact on promoting phosphatase activity. With this approach we simulated ensembles of CaM/calcineurin-regulatory domain complexes using the structure prediction tool, Rosetta, to predict regulatory domain conformations, modeled the extent to which these ensembles occlude the calcineurin catalytic domain through molecular and Brownian dynamics (NAMD and BrownDye), and incorporated the simulated parameters into a Markov model of calcineurin activation (see Fig. 2).<sup>2</sup>

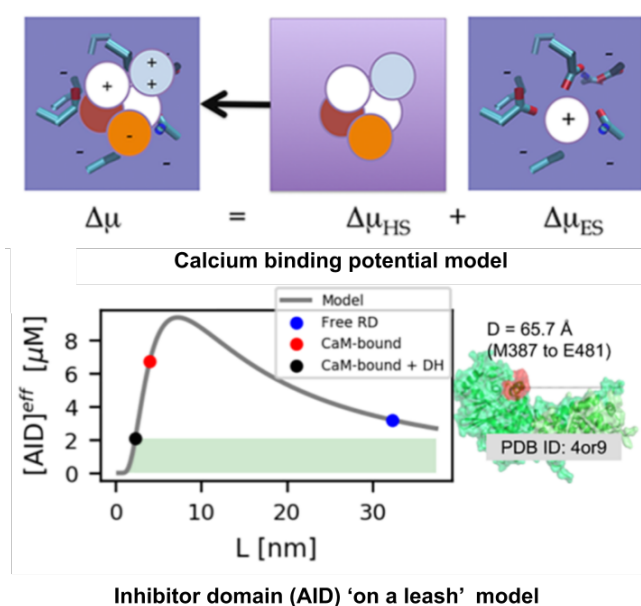


Figure 1: We simulate enzyme thermodynamics using quantities estimated from chemical physics models, including Mean Sphere Approximation theory (top) and polymer physics (bottom) with structural information from molecular simulations.<sup>1-3</sup>

The model confirmed our hypotheses and revealed several interactions between CaM and calcineurin that could expose the catalytic site.

We were excited to validate the importance of these interactions by demonstrating reduced substrate binding affinity in phosphatase assays we conducted with CaM variants [2]; these variants reflected amino acid substitutions that we predicted would weaken CaM binding to the calcineurin regulatory domain. We are now extending these studies using supervised machine learning techniques including Bayesian classifiers and support vector machines using the python sci-kit learn package. These will be used to infer additional molecular attributes shaping CaM' calcium affinity and effects on calcineurin phosphatase activity using data collected from simulation and published experiments. This information will provide a direly-needed guide to engineer calcineurin/CaM binding to precisely control when and how gene expression is initiated by calcium fluctuations in activated macrophages.

## Theme 2: Understanding how intercellular junctions influence ligand signaling.

Calcium binding proteins respond to fluctuations in intracellular calcium to orchestrate calcium-dependent signaling. These fluctuations are frequently triggered by the binding of extracellular molecules, like adenosine triphosphate (ATP), to transmembrane receptors.

*In vitro*, transmembrane receptors can be triggered by pipetting nucleotides into the cellular medium. Triggering of receptors *in vivo* is more complex as the available nucleotide substrate depends on many factors including how cells are packed, diffusion rates through intercellular junctions formed between macrophages and target cells (Fig. 3) and the activity of nucleotidases that synthesize or consume these substrates. Hence, we **hypothesized** that the priming of intracellular calcium signaling pathways in macrophages by extracellular ATP molecules critically depends on the intercellular junction composition, such as the relative distribution of transmembrane receptors to enzymes that metabolize nucleotides. Unfortunately, these factors are difficult to precisely process, monitor or control in experiments, especially when many processes must be observed simultaneously. An appropriately trained computational model in principle could predict how these factors essentially control how a signal, ATP, propagates to junctional receptors *in vivo* to invoke calcium fluctuations. However, most approaches rely on either structurally-realistic atomistic models that cannot be extended to spatial scales (microns) relevant to substrate diffusion or use coarse-grained or continuum models that lack the atomistic accuracy to capture important, atomistic-scale interactions like electrostatic forces between molecules. A key innovation our lab has developed to address this limitation is a **multi-scale approach that models signal transduction as reaction-diffusion processes occurring over large micron-sized regions, using effective parameters estimated from much smaller sub-regions (Fig. 5B).**

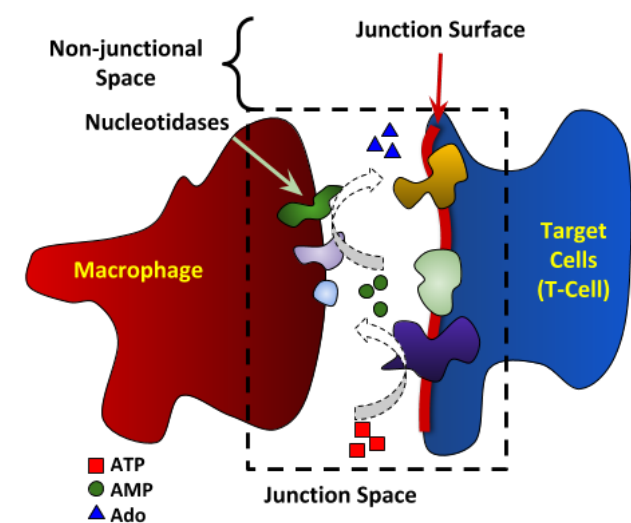


Figure 3: We use molecular simulations and partial differential equation modeling toward understanding complex intercellular signal transduction within macrophage-target cell junctions.

Until recently, a limitation of our approach was that model parameters at the microscopic scale were obtained from the literature or from models that neglected atomistic information. Hence, the parameters we used could not account for subtle changes in surface chemistry, such as pH, or even protein charge distribution, which can significantly impact diffusion and reaction rates. My laboratory has worked toward extending our computational tools for homogenizing interactions between

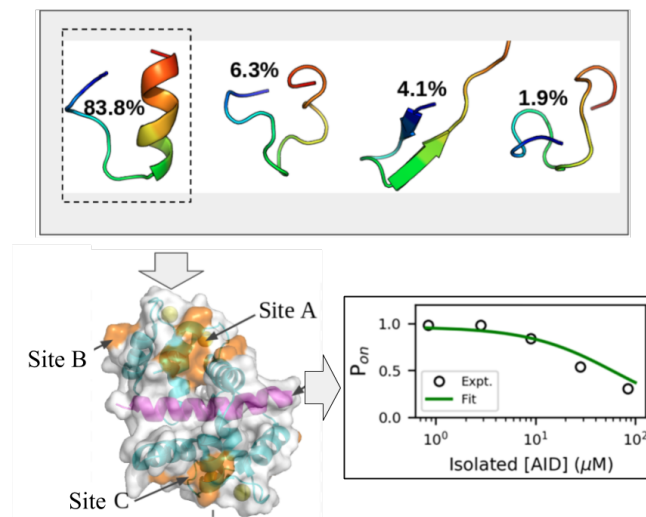


Figure 2: We are developing molecular simulation protocols to ‘dock’ regulatory domain fragments to calmodulin (CaM), from which we can infer native docking poses between calcineurin and CaM and their impact on enzyme activity (right, see also Fig. 1).

Our approach relies on ‘two-scale homogenization’, which obtains ‘macroscopic scale’ properties of solutions to reaction-diffusion partial differential equation (PDE)s, assuming that the PDE parameters can be estimated by solving a related (homogenized) PDE on a smaller scale (Fig. 5B). For biological systems, we make the assumption that the properties of prominent biomolecules, like the distribution and charge of proteins and chromosomes, are conserved across sub-micron regions over which substrate signaling could occur. We then solve a homogenized reaction-diffusion problem using detailed, molecular-resolution structures of those solvated biomolecules in  $100 \times 100 \times 100$  nanometer regions we denote as the ‘unit cell’ (see Fig. 5B). From this modified problem we obtain parameters such as an effective diffusion coefficient that reflects molecular interactions between the diffuser, like ATP, and the solvated molecules at the ‘microscopic’ scale. My laboratory has applied this homogenization approach for a variety of diffusion models,<sup>6–8</sup> including a specialized diffusion model (Smoluchowski) that permits us to include ‘mean field’ potentials arising from charge/membrane interactions. This innovation has for the first time allowed us to probe how the organization of highly-charged enzymes and receptors in intercellular junctions can substantially effect the availability of signaling molecules relative to bulk solution.<sup>9</sup>

diffusers and proteins<sup>6-8</sup> to include sophisticated chemical physics and atomistic scale structural information derived directly from molecular simulation. Our approach to this problem is rather simple, but effective: we use our recent homogenization of the Smoluchowski model<sup>8</sup> that includes 'mean field' potentials and spatially-dependent diffusion coefficients. These quantities are estimated using molecular dynamics simulations that include membranes and receptors.

### Theme 3: Understanding calcium signaling in macrophages

Inflammatory responses in macrophages critically rely on the synchronized activity of several calcium dependent signaling pathways. Given that proteins in these pathways respond to fluctuations of the cytosolic calcium content, changes in the delicate balance of extracellular entry of calcium via channels, for instance, and its extrusion by pumps and exchangers, can modulate pathway activation and ultimately cell function. An example of this behavior that we recently investigated includes how amylin, an insulin-associated peptide that is up-regulated in pre-diabetes, correlated with cardiac dysfunction. It was established that amylin up-regulation intensified intracellular calcium concentration spikes in cardiac cells during contraction, but the mechanism of this effect and how this correlated with cardiac dysfunction was largely unexplained. Since amylin is similar to amyloid beta peptide in its ability to perforate cell membranes, we proposed that amylin disrupted intracellular signaling by permitting unregulated entry of extracellular calcium. **My lab therefore adapted a systems model of cardiac calcium signaling to test whether unregulated calcium entry exhibited in amylin-overexpressing rats was the mechanism behind observed amplifications of calcium fluctuations.**

The system biology model was formed from a large system of coupled differential equations representing the activity of voltage-gated ion channels, pumps, exchangers, ion-binding proteins and diffusion that collectively determine the intracellular calcium concentration. We adapted this model to reproduce experimentally-measured electrophysiology data and calcium fluctuations in amylin-expressing rat cardiomyocytes. After numerically solving this model with PYTHON, we discovered that an amylin-dependent increase in extracellular calcium entry elevated intracellular calcium store content and sensitized intracellular calcium release,<sup>10</sup> both factors contributed to the intensified spikes in amylin over-expressing rats. Surprisingly, we also found that the elevated calcium transients were sufficient to promote CaM-dependent activation of calcineurin and subsequently the transcription factor nuclear factor of activated T-cells, which is a hallmark step in cardiac hypertrophy observed experimentally.

More recently, we have been advancing our system biology modeling to probe similar calcium signaling mechanisms in macrophages, albeit those connected with inflammation (See Fig. 5C).

However, much of our knowledge about macrophages calcium signaling results from rather non-physiological manipulations including gene knock-out or knock-in, or treatments with strong inhibitors, that confound *in vivo* calcium signaling mechanisms underlying macrophage inflammatory responses. While similar challenges that were once significant for cardiomyocytes have been surmounted through years of computational systems modeling, comparable models for macrophages have lagged far behind. Given the breadth of experimental studies implicating calcium in stages of macrophage inflammatory responses, we **hypothesized** that ATP-dependent activation of P2X channels, a class of transmembrane receptors, could invoke cytokine synthesis purely through triggering a small number of calcium dependent signaling pathways. **We challenged this hypothesis by creating the first computational model of microglia,<sup>11</sup> the resident macrophages of the central nervous system.** Our model utilized successful formulations adapted from our cardiac models.<sup>7,12-15</sup> Uniquely, this model aggregated decades of experimental calcium fluorescence, protein expression, mRNA synthesis and channel electrophysiology data specific to macrophages in order to quantify cytokine (tumor necrosis factor alpha, TNF $\alpha$ ) production following ATP exposure. With this model, we predicted the conditions under which physiological ATP release events, such as from within cellular gap junctions, could promote TNF $\alpha$  mRNA production as part of the macrophage's canonical inflammatory response. Interestingly, we also discovered in the development of the model that calcium binding to the CaM/calcineurin complex is a requisite for TNF $\alpha$  production, which could provide another target for engineering CaM constructs we planned to develop.

Nonetheless, a prominent limitation of our model was that it heavily relied on *in vitro* experimental data that did not account for phenotype-specific variations in gene expression and cell morphology observed *in vivo*. We speculated that distinctive changes in macrophage shape, such as ramified (branched) versus amoeboid, that correlate with phenotype would exhibit significant changes in proteins involved in the calcium-dependent pathways we currently model. We are therefore developing sophisticated computer vision tools to test this hypothesis by rapidly characterizing cell morphology as a predictor of phenotype from *in situ* imaging data, using image processing techniques we recently developed for analyzing cardiac cells<sup>16,17</sup> and materials<sup>18</sup>

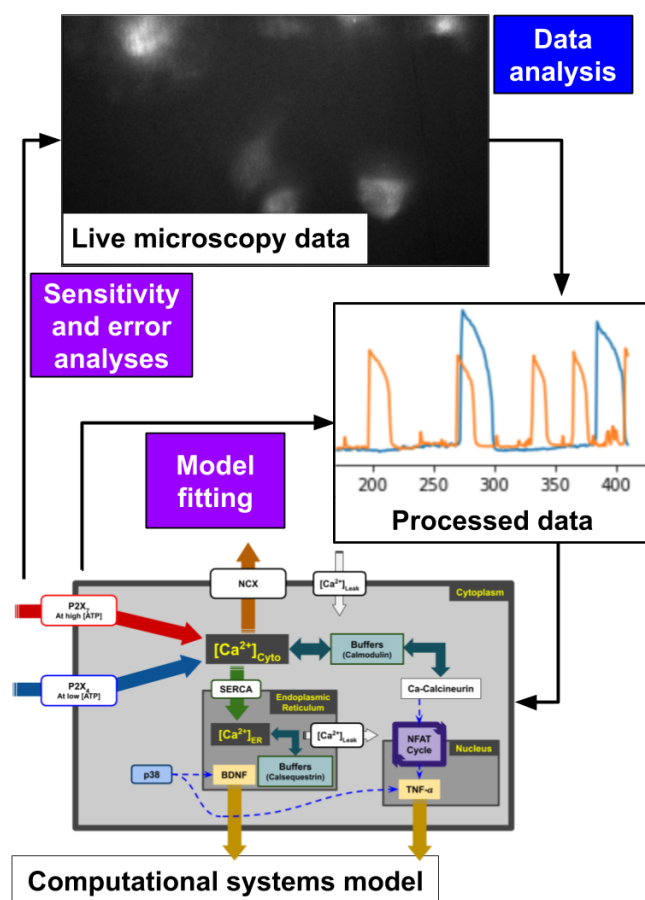


Figure 4: Our high-throughput systems model training uses live cell imaging, computer vision-based data analysis software and a genetic fitting algorithm developed by our lab to fit and test systems models of macrophage signaling.

(see our workflow in Fig. 4). We intend to follow these image classifications with mRNA quantification as a first step towards estimating the expression levels of proteins we consider in our systems model. Importantly, these data can provide cell- and environment-specific parameters for spatially-realistic models of calcium signaling in macrophages as shown in Fig. 5B to predict the production of pro-inflammatory cytokines *in vivo*.

- [1] A. N. Kucharski et al. In: *The Journal of Physical Chemistry B* 120.33 (Aug. 2016), pp. 8617–8630.
- [2] B. Sun et al. In: *Biochemistry* 58.39 (Oct. 2019), pp. 4070–4085.
- [3] B. Sun et al. In: *J. Chem. Theory Comput.* 15.4 (Apr. 2019), pp. 2692–2705.
- [4] B. Sun et al. In: *Front. Mol. Biosci.* (2020).
- [5] B. Sun et al. In: *Biochimica et Biophysica Acta (BBA) - General Subjects* 1862.12 (Dec. 2018), pp. 2651–2659.
- [6] P. M. Kekenes-Huskey et al. In: *The Journal of Physical Chemistry B* 120.33 (Aug. 2016), pp. 8696–8706.
- [7] P. Kekenes-Huskey et al. In: *Biophysical Journal* 105.9 (2013), pp. 2130–2140.
- [8] P. M. Kekenes-Huskey et al. In: *The Journal of chemical physics* 140.17 (May 2014), p. 174106. arXiv: [PMC3535444](https://arxiv.org/abs/PMC3535444).
- [9] H. Rahmaninejad et al. In: *PLoS Comput Biol* (in print) (2020).
- [10] B. D. Stewart et al. In: *Cell Calcium* 71 (May 2018), pp. 65–74. arXiv: [1704.03353](https://arxiv.org/abs/1704.03353).
- [11] B. J. Chun et al. In: *The Journal of Physiology* 597.3 (Feb. 2019), pp. 799–818.
- [12] Y. Cheng et al. In: *Computational Science and Discovery* 5.1 (2012).
- [13] P. M. Kekenes-Huskey et al. In: *Frontiers in Physiology* 3 (2012), p. 351. arXiv: [PMC3463892](https://arxiv.org/abs/PMC3463892).
- [14] J. Hake et al. In: *Journal of Physiology* 590.18 (2012), pp. 4403–4422.
- [15] P. M. Kekenes-Huskey et al. In: *Computational science & discovery* 5.1 (2012), pp. 0–20.
- [16] X. Shen et al. In: *J. Physiol.* 597.2 (Nov. 2019), pp. 399–418.
- [17] D. F. Colli et al. In: *Biophysical Journal* 116.8 (Apr. 2019), pp. 1386–1393.
- [18] B. Sun et al. In: (Oct. 2017).

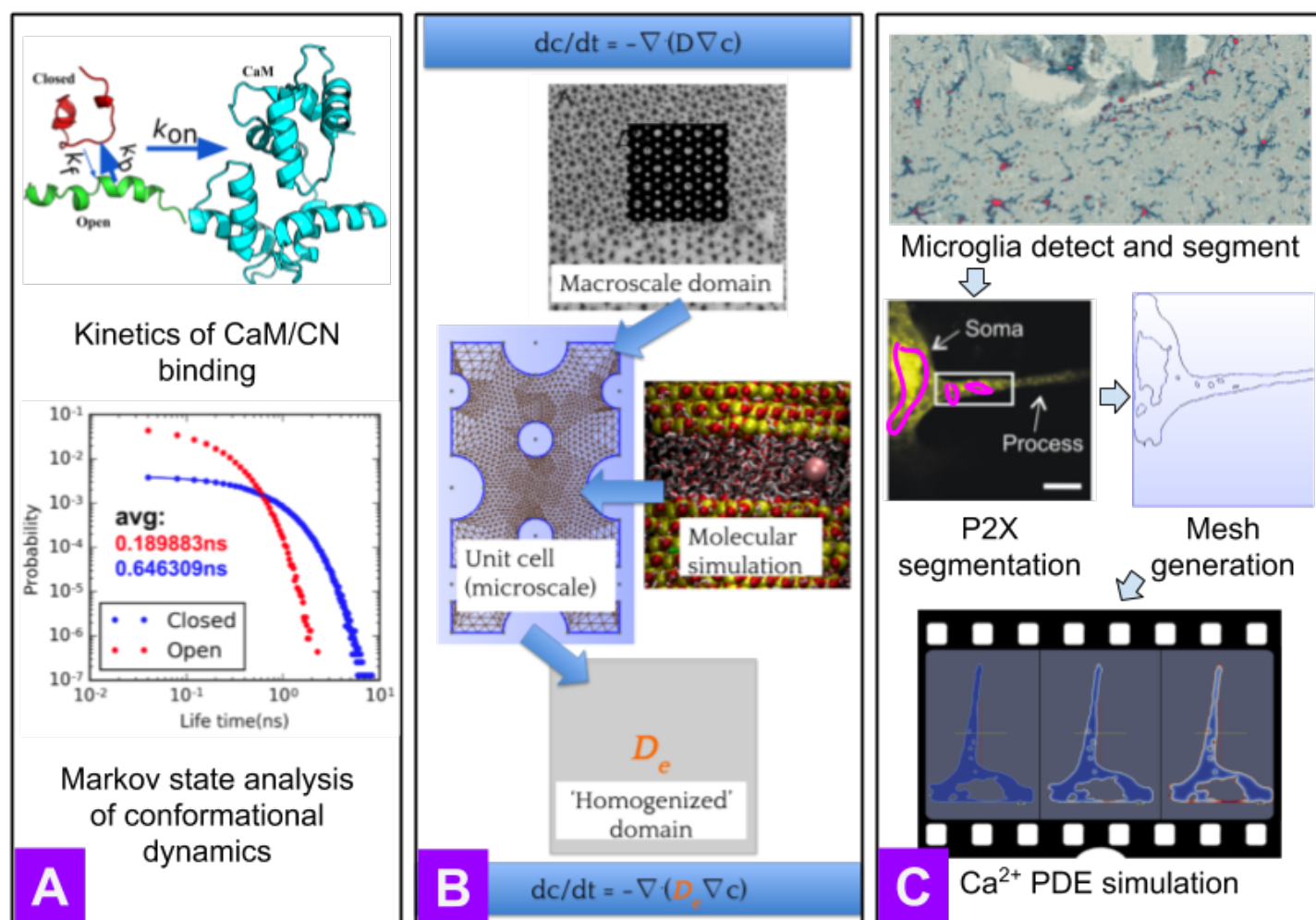


Figure 5: A) Our molecular simulations of calmodulin (CaM)/calcineurin association reveal how disordered domains shape protein-protein binding kinetics;<sup>1</sup> our planned studies will investigate secondary interactions that we believe are critical to calcineurin enzyme activity.<sup>2</sup> B) Our atomistic resolution simulations of small molecule diffusion combined with two-scale homogenization theory have developed will allow us to model second-messenger signaling *in vivo*.<sup>3</sup> C) We are developing systems level and spatially-detailed simulations of purinergic receptor activation<sup>4</sup> to understand macrophage function *in vivo*, using computer-vision techniques our lab has developed.<sup>5,6</sup>

## 0.1 Teaching Philosophy and Innovations

As an educator, I am eager to cultivate excitement and curiosity among my students, while enabling them with the mastery and curiosity to apply scientific principles to problem solving in the chemical sciences. Recognizing that students learn draw from a spectrum of personalities, learning styles and levels of preparedness,<sup>1</sup> my classroom engagement capitalizes on several innovative modalities of communication<sup>2</sup> and implementation. Key innovations include a tiered workshop series for promoting STEM concepts, group-driven problem solving using simple computational methods, and interactive in-class activities using PYTHON, a programming language commonly used in the scientific community.

My dedication to education is most strongly exemplified in the STEM workshops that were in part co-developed by a colleague at the University of Kentucky. We launched our [Mathematics of Physical Chemistry Bootcamp](#) in the Fall of 2015, with the goal of teaching key mathematical and numerical approaches that incoming graduate students would encounter in classes and research. Our **novel angle to the workshop was combining traditional powerpoint lectures with *in silico* laboratories that allowed students to make, test and break simulations<sup>3</sup> appropriate for the covered topics.** The inclusion of a simulation component benefits students by helping students develop visual representations of a chemical species, for instance, to illustrate important physical concepts.<sup>4,5</sup> Survey data following our first offering indicated that students enjoyed the detailed mathematical presentations and the reinforcing tutorials. Self-assessments of their knowledge of the bootcamp topics strongly demonstrated a shift from an average of 2/5 points ('little knowledge') to 4/5 ('considerable knowledge'). Surprising, however, was that nearly one-third of the attendees were undergraduate students, many of which wanted to sharpen their teeth on physical chemistry math before taking the course in the upcoming Fall term. As a result of these data, I have focused on aligning the bootcamp material with the most hard-to-grasp concepts from my physical chemistry courses. Even for incoming graduate students that have had physical chemistry, the incorporation of the simulation labs provides a unique and effective learning experience. Now that my lab has transitioned to Loyola University Chicago (LUC), I am working toward offering a related program to undergraduates and graduates this upcoming academic year.

One of the most difficult struggles I have found in teaching is how to best accommodate the diverse levels of preparedness reflected in the incoming students. In the entry level General Chemistry courses, the disparities were often most apparent, as a significant number of students struggled with the mathematics required for success in the course. Having to fight two fronts, that is, both the mathematics and the chemistry course content, caused many of those students to struggle; unfortunately, when those barriers prove insurmountable, those students are unlikely to advance toward many of UK's degrees in health and the sciences. To address this challenge, building from my positive experiences with the Mathematical Bootcamp, in the Fall of 2016 I created Career Readiness Education in Science and Technology ([CREST](#)) program to engage rising high school juniors in the STEM fields with exercises in computer simulation and programming. I specifically targeted high school students in Eastern Kentucky, a region of the country plagued by high levels of poverty and unemployment. In these monthly workshops, I frequently traveled to Big Sandy Community College to introduce students to a popular programming language, python, and how to use its libraries to generate and analyze scientific data. **Through my CREST program, disadvantaged high school students built competency in computer programming, namely PYTHON, and rigorous quantitative analysis that is rare among entering STEM undergraduates.** Since the original offering of this experience in 2016, I have substantially expanded the scope of the course, including providing more 'jupyter-notebook' based examples that students can run on their own laptop. While CREST was not offered this year due to my lab's move, I am restarting this program this summer.

I have additionally experimented with multimodal communication<sup>2</sup> and cooperative learning<sup>6</sup> teaching methodologies in my traditional courses, including General Chemistry (CHE 105), Physical Chemistry Laboratory, and Physical Chemistry for Engineers (CHE446). In my physical chemistry course, **I actively integrate student-led computational examples to guide learners through difficult concepts.** For example, in one class, I provided students with a jupyter notebook that asked students to enumerating rotational states of a small, diatomic model. Over the ensuing ten minutes, students populated the spreadsheets, plotted the data to analyze graphical trends, and were asked to relate their findings to the analytical approximations for the rotational partition functions we discussed in class. As a result, I found that students performed quite well on the exam question pertaining to statistic mechanical partition functions, therefore these hands-on examples appear help students digest and recapitulate difficult conceptual material.: With each year I offered this course, I continued to expand the number of computational examples provided in supplement to the class. Most recently, I have converted these examples to jupyter-notebooks, which allow students to interactively parameters, simulate and analyze data in ways that were intractable with spreadsheets I had used previously. I am actively developing this approach in my UNIV 102 course at Loyola to expose students to how simulations of chemical processes contribute to scientific research. I am excited about this opportunity to work with more junior students that can directly apply the learned techniques to STEM courses as part of their degree requirements.

A common frustration I hear from students is that the scope of course material is dauntingly broad in chemistry. As a result, progressively synthesizing more advanced ideas by building off of and drawing connections between concepts



Figure 6: Teaching simulation and biophysics through the CREST program for Eastern Kentucky students.

can be challenging. To reduce the barriers to unifying material concepts and orient students in the learning process, I introduce conceptual pictures or 'contact maps'<sup>3</sup> early in course. As one such example, I used a picture of gaseous particles in a box to anchor students around a core molecular perspective from which to understand several chapters' worth of material, including thermodynamics, gas laws and kinetic molecular theory. By explaining material from the perspective of a particle, one could rationalize with simple terms how increasing the temperature would cause molecules to move faster and hit the box walls at a higher velocity, thereby yielding higher pressure ( $pressure = force/area$ ). Distinguishing ideal versus real gas behavior fell naturally from this model, since students could imagine how interactions between particles, such as collisions and long-range attractions, could alter the momentum available to generate an impulse force on the box walls. In other words, *creating a physically-reasonable picture enabled students to predict new phenomena from existing knowledge without formally introducing 'facts' through lecture!* Analogously, I developed a 'central' equation map that related myriad applied formulae we used back to fundamental relationships, such as pressure from  $force = mass * acceleration$ . While my students have not commented directly on this teaching strategy, my technique has been well received. In fact, my semester long assignment in my CHE 446 course require students pair at least two such conceptual models into a simulation of real-life chemical phenomena that is the basis for a semester project. An example from my current class includes a group using a statistical-mechanical model of adsorbate/surface binding to predict adsorption isotherms based on an extended Langmuir adsorption model. It has been very enjoyable to read updates from groups that cover conceptual material I have not yet covered in class!

With this coming academic year, I intend to complement explore machine learning techniques like Natural Language Processing,<sup>7</sup> based on mined text from de-identified student correspondence data, in order to provide models of topic-specific performance based on search strings, inquiries and other tabulated data. In addition, I will emphasize in course, understand patterns in problems that present significant difficulty to students and as a result, improve student outcomes through better-informed teaching practices. We have used one such approach, Bayesian classifiers, to develop predictive models of protein/ion binding affinity based on diverse inputs including protein charge, sequence identity and conservation. Much like traditional statistical approaches, such models weakly depend on the type of input data and output metrics provided are therefore extendable to nearly arbitrary applications, thus we anticipate extension to classroom data will be trivial. This direction is in the very early stages of planning, thus I will seek out advice on maintaining FERPA compliance in such analyses.

Beyond the classroom, I have explored innovative strategies to mentor students in scientific research through my CHE 395 and 790 research courses. I have for instance developed a comprehensive group wiki, the contents of which includes tutorials for a wide variety of our tools, an online repository of our group presentations and manuscripts, as well as an evolving book that explains the theories used in our simulations. This ready availability of research and communication materials has accelerated student productivity, as we have several accepted and published manuscripts from undergraduate and graduate lab members from Chemistry, Physics and Chemical Engineering<sup>8-10</sup> In complement to these materials, I have integrated an online collaboration tool, Slack (Searchable Log of All Conversation and Knowledge), to promote exchanges between all lab members to participate in data discussion and troubleshooting. **My dedication to mentoring students has translated to several awards: 'Teacher Who Made a Difference' (2016), "A&S Award for Innovative Teaching" (2017), and the "Faculty Member of the Week" (2018)**

As I continue learning as an educator and mentor, I will continue to identify opportunities to apply innovative, research-supported best practices in teaching to make learning fun and improve student outcomes. For instance, I am continuing to update my lecture materials to reflect modern applications for concepts I teach in class, as well as increase the number of hands-on simulation activities to foster interactive learning. Much of these efforts can be guided by outcomes obtained in my proposed research, especially since many of the simulations can be reproduced with limited computational resources.

### 0.1.1 Statement on diversity

I believe that success in education stems from a conscientious effort to interact with young students from a variety of cultural and economic backgrounds. As such, I am thankful for the opportunity to work with historically underrepresented groups through my community engagement (high schools and college-level). Since 2012, I have introduced budding young scientists to research in chemistry and how they can prepare themselves for a career in science. I have given talks and informal labs to students at high schools in Carlsbad, downtown San Diego and Chula Vista and Pikesville, Kentucky.

I continue my efforts to reach underrepresented and nontraditional students. I believe that a multi-faceted approach is important to engaging, supporting and retaining such students. Conducting on-campus tours and lab activities, as well as visits to classrooms, are good mechanisms for showing middle-school and high school students that post-secondary education in science is a reality, not just a nebulous aspiration. I am further strongly contributed to building and maintaining a diverse research group. With no hesitation, beyond being a moral objective, I believe a diverse lab, department and university

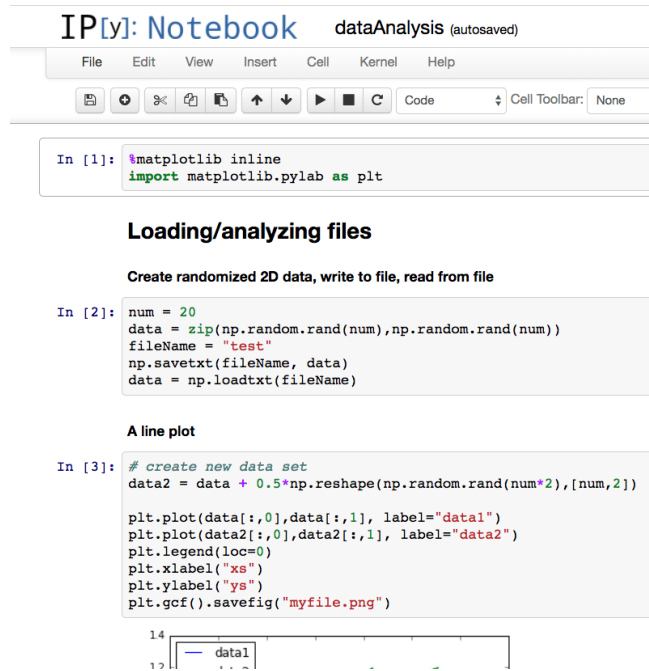


Figure 7: I develop jupyter notebooks that help students understand chemistry through simulation. The notebooks can be written and run in students' web browsers.

Ultimately, this will help me prioritize which types of material to emphasize in course, understand patterns in problems that present significant difficulty to students and as a result, improve student outcomes through better-informed teaching practices. We have used one such approach, Bayesian classifiers, to develop predictive models of protein/ion binding affinity based on diverse inputs including protein charge, sequence identity and conservation. Much like traditional statistical approaches, such models weakly depend on the type of input data and output metrics provided are therefore extendable to nearly arbitrary applications, thus we anticipate extension to classroom data will be trivial. This direction is in the very early stages of planning, thus I will seek out advice on maintaining FERPA compliance in such analyses.

environment are catalysts to new ideas and discovery. My commitment is perhaps most evident in my mentoring, as I have included scholars that are women, those identifying as being LGBTQ, students underrepresented in the physical scientists and scholars from under-served communities (Appalachia).

- [1] A. G. Picciano. In: Journal of the Research Center for Educational Technology 5(1).1 (2009), pp. 4–14.
- [2] C. Jewitt et al. In: Educational Review 53.1 (2001), pp. 5–18.
- [3] P. J. Beuning et al. Research Corporation, 2014.
- [4] E. B. Moore et al. In: Journal of Chemical Education 91.8 (Aug. 2014), pp. 1191–1197.
- [5] H. Tang et al. In: Journal of Chemical Education 93.1 (Jan. 2016), pp. 31–38.
- [6] R. M. Felder et al. In: Aug. 2007, pp. 34–53.
- [7] D. Litman. In: 30th AAAI Conference on Artificial Intelligence, AAAI 2016. 2016.
- [8] B. D. Stewart et al. In: Cell Calcium 71 (May 2018), pp. 65–74. arXiv: [1704.03353](https://arxiv.org/abs/1704.03353).
- [9] B. Sun et al. In: (Oct. 2017).
- [10] A. N. Kucharski et al. In: The Journal of Physical Chemistry B 120.33 (Aug. 2016), pp. 8617–8630.