**Personal Statement**

The interplay of science and engineering fascinated me from an early age. The general theories and clarity of science ground the engineer, and the engineer's optimized inventions and drive to solve society's problems bolster the scientist. I first observed this interplay when I studied the greenhouse effect for a science project in middle school. I found that I could use the greenhouse effect to trap thermal energy from the sun in between a window and a heat-reflecting shade. As the air between them was heated, it would rise to the top of the window and enter the room via natural convection, ultimately heating the room. This scientific study guided me in engineering a prototype the following year to test in my home. The experience of being motivated by a problem, studying it scientifically, and using what I learned to engineer a final product first showed me how science and engineering can work together to benefit society.

While taking science and engineering courses in college, my appreciation of their inter-dependence grew. Consequently, I majored in Engineering Physics, taking pure science courses while applying what I learned to engineering research and class projects. **With deep exposure to both disciplines, I learned the values and language of each, which I am using to bridge the gaps between science and engineering as a researcher and, ultimately, as a professor.**

**Intellectual Merit – Relevant Background**

In my sophomore year, I united science and engineering as an undergraduate researcher in the Fuller Lab in the Chemical Engineering Department at Stanford. **In the Fuller Lab, I led a project to engineer an environmentally friendly process for cleaning silicon wafers by studying the physics of an instability in the flow of water over a rotating wafer.** I loved that I could delve into fundamental fluid dynamics—even learning theoretical methods like linear instability analysis—while applying what I learned to the practical problem of reducing hazardous chemical use in the wafer-processing industry. Although I collaborated with a post doc and a PhD student, I led experimentation, ultimately training other researchers on the apparatus I automated for future projects. In addition to making recommendations for improving this "green" wafer-cleaning process to industry, **I am submitting our findings in a first-author paper to *Physics of Fluids* and a second-author paper to *Physical Review Letters* this year to share our discoveries about the physics of unstable flows on rotating surfaces.**

To compare academic and professional innovation, **I submitted a prototype of my patented energy-saving window to a startup competition at Stanford in my junior year**. Competing with ventures run by Stanford alums and backed by big-name venture capitalists, I progressed through three of four rounds, which gave me opportunities to discuss how to turn my prototype into a product with patent attorneys, venture capitalists, and business leaders. These discussions showed me that I would need to focus on securing investors and exploring potential markets rather than further scientific study or engineering an improved design. Although I considered an offer for Stanford funding to develop my product into a startup, **I realized that my passion for the project came from learning the science and engineering the product**, not creating a business. Instead, I chose to explore academic research further through the Summer Undergraduate Research Fellowships (SURF) program at Caltech.

**As a SURF student at Caltech, I worked in the Troian group to apply numerical simulations to design real-world technology.** Specifically, I studied how an electric field can produce a jet of metal ions for ion-jet propulsion using finite element method (FEM) models in collaboration with NASA's Jet Propulsion Laboratory (JPL). I honed several numerical techniques including scaling analysis, time step selection, refining meshes to show convergence, and benchmarking against previous results. I valued the precise quantification provided by

numerical simulations, but my work felt incomplete without experimental validation. I learned that, like science and engineering, theory and experiment rely on each other, so I sought research in my PhD that would allow me to work closely with theorists.

**At Caltech I am investigating the physics of bubble nucleation in polymer foaming with my co-advisors, Prof. Julie Kornfield and Prof. Zhen-Gang Wang**. **As lead researcher on the experimental portion of this project,** I am responsible for constructing a novel apparatus to perform unprecedented experimental measurements of bubble nucleation. My concurrent study of the theory of foaming will help me to target my experiments on regimes that we predict will exhibit interesting phenomena. The applications of this project to improved thermal insulation also motivate me because they promote energy conservation, the same goal that originally sparked my interest in science and engineering.

**I chose Caltech for my PhD because Caltech, and especially Caltech Chemical Engineering, exemplifies the attitude that science and engineering are best done hand in hand.** Professors are engineering solutions to the world's biggest problems, yet—like me—they value understanding the problems scientifically as much as solving them. Both of my advisors have a knack for identifying fundamental scientific studies with high-impact applications, a skill I hope to cultivate while working with them to design my own research projects in the future.

**Broader Impacts**

One of my favorite parts of science is sharing what I learn with others. I believe that everybody has an intrinsic scientific curiosity, and a good science demo always helps to awaken it. **In college, I sparked curiosity in middle-school students through fluid dynamics demos I built for "Splash,"** a weekend of classes created by Stanford students for local middle- and high-school students. My favorite demo challenged students to discover the connection between the flow of electricity in a wire and the flow of water in a tube by using equipment I provided to design and run their own experiments. **Splash works to make the program accessible to students of diverse backgrounds,** so naturally some students felt less comfortable speaking up. Nevertheless, after they saw their hypotheses either supported or refuted by the first experiment, *all* the students—shy, aloof, and outgoing—became invested in coming up with the "right" hypothesis and proposing the next experiment.

**At Caltech, I have continued making fluid dynamics demos to share with the public. In one demo, I used equipment I built for my research to introduce passers-by to microfluidic devices at Caltech's science awareness day, "Science for March."** In the demo, I took two jugs of water dyed different colors and showed that, while they mix when poured into a cup, in a microscopic channel they flow alongside each other without mixing. As perplexed passers-by discussed their hypotheses for this result of turbulent vs. laminar flow, I shared how this simple concept has been used to produce microfluidic devices that can diagnose HIV and simulate human organs. Since then, **I have convinced friends to join me in sharing my demo at the 626 Night Market,** a local attraction attracting tens of thousands of attendees. It is important to me to show the public how scientific concepts fuel technological development so they will feel more invested in supporting science research in the future.

I serve the local community in other ways besides science demos. **For almost a year, I have been tutoring high school students** from Pasadena's public school system in math and science. **The program focuses on serving students from under-represented backgrounds in STEM fields** so, as a tutor, I also serve as a role model for those with less exposure to STEM careers. Additionally, I volunteer with Caltech students to cook dinner at the Pasadena homeless

shelter each month. These experiences allow me to connect with fellow Pasadenans as a community member and as a PhD student.

**I also value sharing science globally**. **Last year, I planned science experiment kits for schools in villages during my summer internship with the Luys Foundation in Armenia.** Though ethnically Armenian, I have little experience with the country of Armenia, so I learned to work with native Armenians to understand local needs. To overcome the lack of resources available to these schools, my co-workers and I based the kits on locally grown crops, such as grapes, apricots, and tomatoes. **To reach a broader global audience, I use the internet for outreach.** Last year, I worked with Collegeprep.org to prepare an ACT tutoring video for students without access to expensive ACT prep guides. When I improve my video-editing, I plan to record my science demos and publish them on a YouTube channel, *Wait, WHAT?*, focused on explaining surprising phenomena in everyday life. I also look forward to writing articles to share my research with the public through "Caltech Letters" this year. Last, having had such a transformative experience as an undergraduate researcher in Caltech's SURF program, **I want to pass on that experience to other undergrads by mentoring a SURF student next summer**. Undergraduate research is the best way to prepare for a PhD program, so **I want to encourage broader participation in PhD programs by providing that opportunity to others.**

**At Caltech, I share my passion for science through teaching.** This fall, I am excited to be the **teaching assistant (TA) for the first-year graduate kinetics course**. In particular, I want to motivate course concepts with demos from my research on the kinetics of polymer foaming reactions. Additionally, **having been a TA as an undergraduate**, I know that a TA can serve as both a source of knowledge and a mentor, and I am enjoying guiding the incoming chemical engineering PhD students through selecting an advisor and managing coursework their first term.

**I am also the new safety officer in the Kornfield Lab.** I plan to implement safety practices that I learn from my work with the Dow Chemical Company, JPL, and other Caltech professors, especially those pertaining to high-pressure fluids and handling hazardous chemicals. **This experience will help me with my future goal of leading a research group,** where I will be responsible for the safe design of my lab and safe practices of my students.

**Future Goals**

Over the course of my PhD, I will seek out opportunities to connect with other scientists and share my research. Next year, I plan to attend the **Society of Rheology Conference** to meet scientists in my field. In 2020, I intend to share results from my research at the **Gordon Research Conference on polymer physics** to establish myself in the community. Science is inherently international, so I am also looking forward to strengthening connections I have made with professors outside the US. I continually look for overlap between my research and theirs to spark international collaborations so I can experience how science is done outside the US.

**Ultimately, I aspire to work in academia as a research professor because I have found that I am driven most by my scientific curiosity and my passion for sharing that curiosity with others.** I dream of starting my own research program focused on studying the fundamental physics and chemistry of the production of important chemical products, both to improve existing products and to innovate new ones. As a research professor, I would continue to do outreach by sharing demos of my research with the public at events and through online media platforms. Most meaningful to me would be the opportunity to touch the lives of my students, both as a teacher in the classroom and as a mentor in the lab. **The NSF GRFP would allow me to focus on gaining the scientific, mentorship, and leadership experience I will need to make a strong impact on science, scientists, and the public as a professor.**

## Intellectual Merit – Motivation and Background

Foams have the highest thermal resistance and impact strength per weight of mass-produced materials. Research has shown that pushing these boundaries to produce unprecedented thermal insulation and impact-resistance materials requires nucleation of more bubbles during foaming[1]. Current methods of foaming have reached limits on the number of bubbles they can nucleate before they coalesce, but recent theoretical models suggest a new approach. When foaming polymers with $CO_2$ dissolved under sufficient saturation pressure, rather than nucleating directly as a vapor**, bubbles can first nucleate as a *liquid* with a much lower nucleation barrier, *exponentially* increasing the rate of bubble nucleation**[2]. Harnessing this nucleation pathway could lead to materials with previously unattainable thermal and mechanical properties.

Because of the short lifetime and small size of liquid bubbles, they have so far eluded experimental observation[3]. **I propose to make the first observation of liquid bubble nucleation during polymer foaming by constructing a novel apparatus with the appropriate temporal and spatial resolution. I will then characterize both the conditions required for liquid bubble nucleation and the nucleation barrier.**

### (1) Method for Long-term Observation of Nucleation

Liquid bubbles quickly vaporize after nucleating[2], making direct observation challenging. By instead flowing the nucleating fluid at high speeds (~ 1 m/s), I can stretch out the brief lifetime of a liquid bubble along the length of a channel. This technique from microfluidics, "hydrodynamic focusing," permits continuous observation of each moment of a bubble's life. In my adaptation (Fig. 1), a "sheath stream" (A) of pure polymer confines a "central stream" (B) of polymer and dissolved $CO_2$ within the central, constant-velocity region of the Poiseuille velocity profile. Thus, the elapsed time $t$ corresponds to the distance traveled along the channel $x$ divided by the velocity at the center of the flow $v$. At 1 m/s, the segment of the channel probed by a standard 100 μm-wide X-ray beam[5] would correspond to 0.1 ms in the nucleation timeline, **allowing measurements with sub-millisecond resolution**.

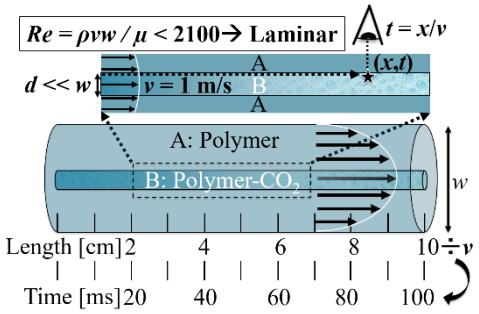



Fig 1: Hydrodynamic focusing. **Top:** close-up of nucleating stream, $t = x / v$. **Middle:** channel. **Bottom:** map of timeline onto length.

The small channel size and high viscosity of the polymers ensure a low enough Reynolds number (*Re*) for laminar flow (Fig. 1, top). They also result in a fast pressure drop (~10s MPa/s at 1 m/s flow), which I can control according to Poiseuille's Law to induce nucleation. Additionally, hydrodynamic focusing prevents heterogeneous nucleation along the walls.

***Preliminary Work*** I have constructed a hydrodynamic-focusing microfluidic device that can withstand the pressures commonly used when foaming polymers with $CO_2$ (~10 MPa)[3]. With image segmentation software adapted from my undergraduate research, **I verified that I can produce a centered inner stream (B) of a desired width.** Using this device, I have observed nucleation of dissolved $N_2$ gas in glycerol with a high-speed camera at pressures up to 0.5 MPa.

**I am collaborating with scientists at NASA JPL, Caltech, and Dow Chemical** to modify this apparatus to control the inlet and outlet pressure within 0.1–10 MPa based on the design of Chen-Jolly *et al*[6]. High pressure is produced with an ISCO syringe pump and high-pressure liquid chromatography (HPLC) pumps, and is controlled with back pressure regulators.

### (2) Observation of Liquid Bubble Nucleation: Small-angle X-ray Scattering (SAXS)

Detecting nanometer-sized bubbles is possible with small-angle X-ray scattering (SAXS), which can detect features from 1–100 nm in amorphous systems[5]. I will first use optical

microscopy to identify the section of the channel appropriate for SAXS by finding where bubbles are too small to see (~100s nm). I will then modify my apparatus to be portable to the DND-CAT beamline at Argonne National Lab, where I will perform SAXS with assistance from experts at Dow Chemical and experienced lab members. I will use scattering data from SAXS to compute the mass density of bubbles[5] and thereby distinguish between liquid and vapor bubbles. To determine whether liquid bubble nucleation has a lower nucleation barrier, I will also measure the onset of nucleation and the nucleation rate, from which I can estimate the nucleation barrier.

**(3) Demonstration of Liquid Bubble Nucleation: Validation of Theoretical Model**

While previous predictions showed that nucleating liquid bubbles requires saturation at pressures exceeding 25 MPa[2], my colleague Dr. Huikuan Chao in the Wang Group at Caltech has used perturbed-chain statistical associating fluid theory (PC-SAFT) to show that **adding a volatile liquid** (cyclopentane) to the vapor-polymer solution ($CO_2$ dissolved in a polyol) **can permit liquid bubbles to nucleate at more accessible saturation pressures (~8 MPa).** I plan to look for signatures of liquid bubbles with SAXS to determine whether adding cyclopentane at the predicted threshold concentration causes liquid bubble nucleation. **Demonstrating liquid bubble nucleation will help to validate existing theories of bubble nucleation in polymer foaming, a crucial step for advancing theoretical models of polymer foaming[3]**.

**(4) Characterization: Conditions, Nucleation Barrier, and Vaporization of Liquid Bubbles**

If I can demonstrate liquid bubble nucleation with SAXS, I will work with Dr. Chao to expand the "phase diagram" of required conditions proposed by Xu *et al.*[2] to include the effect of adding a volatile liquid and verify it experimentally. We will also estimate the corresponding nucleation barriers experimentally and theoretically to lay the foundations for a quantitative model of the nucleation barrier that could guide future foam production. Additionally, I want to investigate the effect of liquid bubble nucleation on the distribution of final bubble size. I hypothesize that liquid bubbles will nucleate within a shorter time than vapor bubbles and will subsequently vaporize at the same pressure, leading to a more monodisperse foam. Currently, there are **no theories for the vaporization of liquid bubbles in polymer foaming**[3], so this study will drive development of new models.

**Broader Impacts**

The lower nucleation barrier of liquid bubble nucleation exponentially increases the number of bubbles that can be nucleated during foaming, which allows the production of foams with better thermal and mechanical properties. For example, **the leading strategy to increase the thermal resistance of industry-leading polyurethane insulating foams is to increase the cell number density[1].** Over 15% of energy consumed in the US is used to control the temperature of buildings, which proper insulation can reduce by 30–50% per building[7]. Because polyurethane has the best thermal resistance on the market, improving its thermal resistance will incentivize installation of proper insulation and encourage higher building efficiency standards.

**With the NSF GRFP, I will be able to focus more on fundamental understanding of how liquid bubbles nucleate**. I will have the freedom to focus on the foaming of simple, model polymers, well-suited to quantitatively characterize this phenomenon, rather than proprietary polymer mixtures relevant to industrial applications. **The results of this study will lay the foundation for new models of polymer foaming and will open the door to a new regime of foamed materials with unprecedented thermal and mechanical properties.**

**1)** Okolieocha, C *et al. Eur. Polym. J.* **2015**, *73*, 500. **2)** Xu, X. *et al. Soft Matter* **2013**, *9*, 9675. **3)** Di Maio, E. and Kiran, E. *J. Supercrit. Fluids* **2018**, *134*, 157. **4)** Ghazal, A., *et al. Lab Chip* **2016**, *16* (22), 4263. **5)** Svergun, D. I. and Koch, M. H. J. *Reports Prog. Phys.* **2003**, *66* (10), 1735. **6)** Chen-Jolly, H. *et al. Chem. Eng. J.* **2018**, *334*, 389. **7)** *Saving Energy and Money With Spray Polyurethane Foam (SPF) Insulation*; Salt Lake City, UT, 2007.

**Introduction:**

The importance of interdisciplinary research has become increasingly clear as academic silos have imposed divergent evolution of thought. Physicists love to simplify systems, biologists revel in the complications of outside interactions, and mathematicians are philosophers interested in the abstract. While computer scientists seek optimal extremes, physicists focus on transitions between them. Working across subjects enables researchers to think outside of conventions, using the creativity in many fields to find new solutions. I chose to study at Caltech because it is a small program thriving on interdisciplinary research.

The proliferation of computational power has only added to the importance of an interdisciplinary approach to research. Unfortunately, the tools of other subjects can easily be misused if their nuances are not fully understood; scientists may choose the most popular tool instead of the right one. Worse still, scientists may be tempted to use computational brute force instead of fully investigating the essence of the problem. In computer science raw computation can be overemphasized at the expense of *learning* from other fields. The most essential interdisciplinary research in my opinion reaches beyond re-purposed machinery and draws from domain-specific perspectives.

Rigor and relevance are the pillars of deeper interdisciplinary research. My research questions should ideally have relevant applications in other areas of science while being addressed rigorously and in mathematically sound ways. These concepts can be difficult to balance: overemphasizing mathematical rigor may oversimplify a system and reduce the usefulness of its applications. Alternatively, improving relevance may come at the expense of domain specific assumptions that lessen the broader impact of the mathematical result. My unique combination of mathematical fluency, creativity, and understanding of applied disciplines will help me balance these trade-offs.

**Intellectual Merit:**

Two areas of applied research I have engaged in are RNA folding and cancer genetics - both of which are on the boundary of computer science and biology. My in depth study of these fields has enabled me to think like scientists from different areas, uniquely positioning me to facilitate a deeper exchange of ideas. Hence, both of my research projects have been based on interdisciplinary *concepts* instead of only applying their *machinery*.

My RNA folding research was at Williams College with Professor Daniel Aalberts. RNA folding algorithms have historically been developed by computer scientists and biologists and been focused on finding a single optimal minimum free energy structure. This is appealing to computer scientists because it provides an elegant dynamic programming algorithm. Biologists appreciate the concept of a "correct" structure as well since it enables them to assign each molecule a function within their larger systems. These approaches, however, do not incorporate a fundamental perspective from statistical physics: entropy. In RNA, any given folding state is very improbable and thermodynamically close to *many other states*. As a result, the field has been developed completely around studying folding states with minuscule probability. While these states provide useful insights into possible folding structures, the energy of the optimal state provides an imperfect predictor of the probabilities of those structures. Our research showed multiple cases where a sub-optimal folding structure became dominant once the probabilistic weight of its fluctuations was included. Our algorithm makes use of the *partition function* to define *macrostates* which contain the probabilistic weights of all the fluctuations of a structure. I implemented this algorithm and developed a method to characterize the kinetics between these macrostates. I tested this kinetic

theory using simulations and have developed a method to move macrostates into a hierarchical trees to represent their recursive structure. Part of this work is in the publication process and I am currently running additional simulations to show the many implications of this new kinetic theory. Our research, which we think will take future work in a new direction, was only possible because of the introduction of concepts from physics.

At Caltech, I am currently working as part of Dr. Bruck's Team trying to predict the risks of certain cancers using specific markers embedded in the genome of non-cancerous cells. DNA is traditionally viewed as a *source* of information, hence biologists search for genes that control specific phenotypes in the body. Our team has backgrounds in many academic areas including electrical engineering which led us to view the genome as a *signal*. Instead of the standard approach of looking for a gene that controlled cancer risk, we searched for signs that the genome is mutating quickly - suggesting that it is prone to becoming cancerous. My contributions have primarily been in finding an accurate classifier for cancer type and displaying the classification results intuitively. In addition to finding a successful classifier, we have investigated the reasons for the classifiers success and carefully examined statistical biases in testing procedures. I have investigated relative success of different classifiers and verified that they agree with the theory we believe we are measuring. Our work makes important contributions to how the genome is viewed; not just as a source of genes, but also as a window to a stochastic process. We will be submitting our paper to *Nature* next month. Our new approach will hopefully be fundamental in studying diseases that are caused by changes in the genome itself. These findings were only possible because of our team's interdisciplinary background.

After graduating from Williams, I sought to build my mathematical foundation and strengthen my capability to do rigorous theoretical research by studying at the University of Cambridge on a Herchel Smith fellowship. I successfully passed the Level 1A mathematics exam while emphasizing courses in mathematics and theoretical computer science that helped build my research background, rather than the courses covered in the exam. Cambridge is well known for its dedication to rigor in mathematics and computer science, making the experience ideal for developing my mathematical maturity. The contacts I built are invaluable to my understanding of a foreign academic system and will hopefully prime future cross-border collaborations.

My applied projects have helped make my fundamental research impactful. A primary concern that surfaced in the Bruck cancer research was avoiding the hidden biases that emerge when working with combined datasets. Properly addressing this led me to define a new problem with Professor Leonard Schulman: developing an algorithm to identify hidden variables in causal graphs. Our goal is to suggest where the hidden variables interfering in any graph are given data on the joint distributions. The connections between these two research problems are additional evidence that rigor and relevance are in fact complements: study of an applied area often leads to relevant problems to investigate rigorously.

**Broader Impact:**

In my applied research, I hope to enforce the understanding of statistical biases. The cancer-prediction classifier that we have developed has been surprisingly, and even suspiciously accurate. Midway through our work, we discovered that some files were marked with a letter "D" and others with the letter "W." Follow up indicated that these data files had been sequenced using a slightly different technology. Although the curators of the database believed the difference was insignificant, our classifier distinguished "D" from "W" files with $98\%$ accuracy. Obviously if a set of

cancer patients is mostly D-files and a set of healthy patients is mostly W-files, a D vs. W classifier is going to perform fairly well despite containing little information about whether a new patient is likely to develop cancer. We have posted a summary of our results on the archive *arXiv* in order to inform other scientists of this bias.

I have been making a broader impact at Caltech working as an assistant coach for the Caltech cross country and track teams. Athletics at Caltech strives to build community, with students often finding an irreplaceable support network in their sports teams. In addition, students represent Caltech when they travel to competitions. These goals extend beyond athletic success and I take my role as a mentor and leader of new scientists seriously.

Success in running requires athletes to be disciplined in their training, academic, and personal lives. I mentor students to find creative and productive strategies to manage their time including setting goals as a team. I recommend students pursue depth over breadth in their classes and use their extra time to follow their interests beyond the required material. My students agree that the structure and mentorship they get from the team leads them to healthier and more productive academic lives.

Retention is an increasing problem in STEM because students sometimes quickly loose confidence in the face of difficulty. Similarly, running results are not always immediate; runners must rely on confidence in their abilities and work-ethic for long stretches before benefits are realized. I encourage my athletes to face challenges with confidence and persistence to prime them for success on and off the track.

These lessons carry special weight because of my background as both a graduate student and a semi-professional runner aspiring to compete in the 2020 Olympics (I am currently in the top 20 in the U.S. for the Half Marathon). Having been a successful student-athlete in undergrad who worked his way into graduate study, I have demonstrated ability to follow through on personal goals. This makes me a valuable resource to students, particularly those aspiring to continue to graduate school.

I have also extended my reach beyond coaching. Last year, I wrote an article on social media explaining the dangers of eating disorders in running and why a correlation between slimness and speed should not be confused with causation. I will be publishing this article on a running website to help promote scientific understanding in an important issue. In addition, I have developed a statistical model to normalize cross country races by their difficulties, hilliness and weather, which I used to rank the top teams in the country. I have shared this model publicly and hope its reach can help inspire interest in math and statistics, much like other statistical sports websites.

**Future Goals:**

I have thrived balancing applied and theoretical interdisciplinary projects. Two of these research projects have led to new interesting results where my diverse background played an important role in approaching the problem differently. The support of the NSF GRFP will allow me to further pursue this balanced approach to rigor and relevance on interdisciplinary problems by simultaneously working on an applied project with the Bruck Lab and broad-reaching theory with Dr. Schulman. In addition, I hope to use my status as an NSF Fellow to fill an often missing role in sports as an ambassador for science.

The difference between causation and correlation is a fundamental problem in data science as at times there is little understanding of why a model works or the exact scientific meaning of a relationship. In an article published in *Nature*, researchers found a link between nighttime light and the development of nearsightedness [1]. A later study discovered that this relationship stemmed from nearsighted parents being more likely to leave lights on and also to have nearsighted children [2]. The parents' nearsightedness is an example of a *hidden variable* that induced correlation without causation.

**Research Problem Statement:** In my research with Dr. Leonard Schulman, we begin with a Bayesian network that represents a potentially incomplete causal relationship between variables. We hope to develop an algorithm that identifies unobserved variables from inconsistencies between the data and the hypothesis causal graph. These variables represent the minimal resolution of these discrepancies that still maintains known causal relationships.

**Background:** A causal graph is a Bayesian network that represents directed causal relationships between variables. Judea Pearl has extensively developed his "do-calculus," allowing the calculation of interventional probability distributions [3]. The causal probability $\Pr(y \mid \hat{x})$ differs from regular conditional probability $\Pr(y \mid x)$ in that it measures the probability distribution of $Y$ while *holding* $X = x$ at a given value rather than simply *knowing* its value. An example of a causal graph is provided in Figure 1. In this example, both $X$ and $Y$ may be highly correlated with a third variable, $U$. With conditional probabilities, Bayes rule allows knowledge of $X$ to be converted into information about $U$, which in turn gives information about $Y$. In Pearl's do-calculus, holding $X$ constant and averaging over $U$ severs this connection: $\Pr(y \mid \hat{x}) = \sum_{u \in U} \Pr(u) \Pr(y \mid x, u)$.

Our algorithm will be built on the concept of a *mixture* or the sampling of distributions from a number of sources with given probabilities. For example, a mixture of two distributions can be obtained by flipping a coin chosen from a bag of three silver coins that are biased towards heads and two gold coins that are biased towards tails. The goal when learning a mixture is to obtain both the probability of each source (silver vs. gold) as well as the probability distribution of each source (probability of heads vs. tails given the composition of the coin) from the data samples. Previous work includes algorithms for learning mixtures of product distributions [4].

**Approach**: Learnability of $U$ can best be understood by looking at the example provided in Figure 2. If it can be confidently determined that $X \to Z \to Y$ and there is a stronger correlation between $X$ and $Y$ than the link between either variable and $Z$, there is evidence of a hidden variable affecting $X$ and $Y$. In Figure 1, however, the existence of $U$ cannot generate new distributions. Our algorithm will focus only on learnable situations which we will need to define precisely.

One key insight to our approach is that learning the existence of some unobserved variable $U$ is equivalent to learning a mixture of distributions of the form given by the sub-graph of all the variables affected by $U$. In Figure 2 for example, conditioning on some value $u \in U$ gives a distribution of the form $X \to Z \to Y$. Learning $U$ is equivalent to learning a mixture of distributions on $X \to Z \to Y$ for which all of the distributions behave the same way with respect to $Z$ since $Z$ is not affected by $U$. In the earlier nearsightedness example, the vision of the parents defines which "source" of child and light distributions is received, one of which includes higher probabilities of nighttime light and nearsighted children. Hence, our problem can be solved by learning mixtures of distributions on arbitrary Bayesian networks.

I have developed an iterative algorithm for learning these mixtures. This gives a locally optimal solution and I am currently investigating proofs of accuracy bounds. Iterative methods, however; have known biases and we are investigating a more exact algorithm. Insight for better algorithms

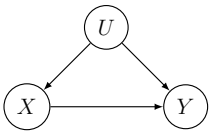


Figure 1

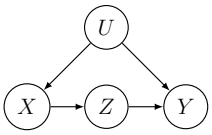


Figure 2

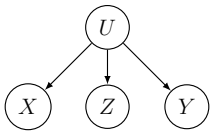


Figure 3

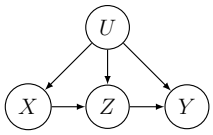


Figure 4

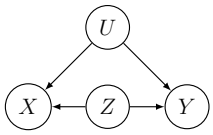


Figure 5

may come from related work on learning mixtures of product distributions which are effectively Bayesian networks with arrows only from $U$ (Figure 3). One particularly interesting solution uses a gridding approach to linearize equations for second moments, solving for the parameters directly [4]. We are working to adapt this solution to learn mixtures of Markov chains, shown in Figure 4. These findings will then be used to develop algorithms to study mixtures of arbitrary Bayesian networks, allowing us to search for a potential hidden variable $U$ on any causal structure.

**Intellectual Merit:** Machine learning classifiers can at times mistakenly learn to predict hidden variables like $U$ instead of the intended $X \rightarrow Y$ relationship. In my cancer research with the Bruck Team for example, we discovered that two DNA sequencing methods left differentiable signals despite being previously believed to be equivalent. Our "predictions" were often overly accurate because the cancer patient and healthy patient data were taken from differing sequencing methods. Our classifier had mistakenly learned the signal of the sequencing method ($U$) instead of the signal between the genome and the development of cancer.

Our cancer project is similar to Figure 1 where $U$ cannot be learned because there are not enough equations of moments to determine the unknowns. While we were able to identify the effect of sequencing techniques once we knew $U$ existed, we would like to be able to find other problems in combined data-sets without having to guess what they are. Insights from our work may show how to set up informative causal sub-graphs, allowing for the detectability of hidden variables. The Bruck Team cancer model could for example look like Figure 5, with the mutation rate ($Z$) determining both the result of our classifier ($X$) and the risk of cancer ($Y$). If $Z$ is measured, it will have a learnable $U$ for the same reasons as Figure 2. Once our algorithm verifies that the relationship between the classifier and cancer is unaffected by hidden variables, we no longer have to measure the mutation rate ($Z$) to make confident predictions.

**Broader Impact:** Identification of hidden variables could help researchers determine if their models are incomplete. A biologist seeking to understand a complex biochemical system may have identified many of the components in a chain but want to investigate where to look for additional catalysts affecting various steps. Similarly, an economist suggesting that changing taxes would prevent recessions could use our work to ensure that the intended outcome is not only correlated with the proposed economic policy, but also causally linked. This work has the potential to transform our confidence in scientific results by providing a rigorous way to verify the completeness of models.

## References

[1] G. E. Quinn, C. H. Shin, M. G. Maguire, and R. A. Stone, "Myopia and ambient lighting at night," *Nature*, vol. 399, no. 6732, p. 113, 1999.

[2] K. Zadnik, L. A. Jones, B. C. Irvin, R. N. Kleinstein, R. E. Manny, J. A. Shin, and D. O. Mutti, "Vision: Myopia and ambient night-time lighting," *Nature*, vol. 404, no. 6774, p. 143, 2000.

[3] J. Pearl, *Causality*. Cambridge university press, 2009.

[4] J. Feldman, R. O'Donnell, and R. Servedio, "Learning mixtures of product distributions over discrete domains," *SIAM Journal on Computing*, vol. 37, no. 5, pp. 1536–1564, 2008.

Rather than being the kid taking apart radios and doing science experiments, I was the kid drawing on the walls, on the floor, and on my sister's face (with permanent markers). I immersed myself fully in the visual arts upon moving from Upstate NY to Israel, at the age of ten, largely due to the language barrier I experienced as an immigrant child not yet fluent in Hebrew. I self-navigated towards the world of art and design for many years, limiting my exposure to STEM, and it was not until college, while majoring in architecture, that I began to consider science as an option. Eventually, I came to realize that research is a sequence of design problems and that scientific methods are a powerful tool for solving these problems. Fast forward 10 years, and now **I am a graduate student at Caltech pursuing a PhD in Civil Engineering**.

**Journey to Graduate School:**

In my last year of undergraduate studies I became particularly excited about space technologies and the urbanization of Mars. When a NASA competition called for ideas to 3D print houses on Mars, I was eager to give it a shot. I recruited a team and we submitted our design phase proposal, competing against ~150 teams from around the world, and won an honorable mention. As we progressed in the competition, I managed a large team of scientists ranging from undergraduates to postdocs. When our budget ran out, we merged with Caltech and I was appointed project manager. I moved from Israel to California to follow my dream of developing new technologies that could provide fast, affordable housing for our world and beyond. This experience taught me that the sky is not the limit, and I became even more determined to develop game-changing building technologies.

Inspired by my teammates at Caltech, I learned that **I wanted to do science, not just manage it**. I began to self-study intensely with both online and on campus classes to prove to myself and to the engineering department at Caltech my academic interest and aptitude. As I learned more about Mars, I further realized how precious Earth is, which stimulated my interest in developing biomaterials that would reduce waste and minimize the impacts of global warming. I am now working on a project in Professor Chiara Daraio's lab to create a new biomaterial from algae and agricultural waste to pursue that vision. I am excited to work on research that will lead to an increased availability of sustainable building materials.

**Intellectual Merit**

Through my architecture studies, I increasingly explored new technologies and attempted to merge these two fields in a meaningful way: I see myself not just as an architect, but as an 'archi-tech', combining architecture and technology.

For example, in one project I developed a new method for crowdsourced urban planning based on Facebook. My project used a new mapping technique to generate "pirate" aerial maps by flying a kite with a camera rig above the site. My maps were then used as evidence in Supreme Court to document a contested highway in East Jerusalem that the municipality pushed through without due process, resulting in revised plans for the road. Another technology driven project I developed created a system for moving buildings that used a scissor lift and conveyor rollers to move building sections in 3-dimensions.

Due to my keen desire to learn more about technology I spent my last year of undergraduate studies working as a Venture Capital tech analyst. In this position I was exposed to exciting new technologies through the business lens of startup investing. I progressed rapidly from analyst to fund manager and was appointed Director of the Jnext Technology fund, a $20M government sponsored economic development program for the Jerusalem Development Authority.

Working in economic development has many commonalities with doing research. When implementing a new economic development policy, the process is often iterative. I would start out with small-scale "experiments" of a new policy idea, and if successfully, I would then scale it up. I also frequently conducted "literature reviews" to see what other cities were doing, and if their policies could be suitable for Jerusalem. Many times projects failed, budgets fell through, or partnerships crumbled–sending me back to the drawing board. There, like in my research, persistence and conviction kept me on track to pursue my goals, despite the challenges. Ultimately, my impactful work at Jnext greatly contributed to an increase of 60% in the number of startup companies in the city and led to an increase of ~6000 jobs over two years through my various activities including: writing new grant policies, providing subsidized office space to companies, recruiting support, and creating partnerships. This rise in Jerusalem-based startups led foreign dignitaries, global innovation managers, and international press to flock to Jerusalem to learn if our "secret sauce" could be replicated successfully in other locations. The global interest in my work led me to co-author an article outlining the economic insights behind Jerusalem's tech scene published in the Coller Venture Capital Journal. My work in Jerusalem was further recognized when in 2017 I was selected for the Forbes Israel 30 Under 30 list and was the youngest member appointed to Jerusalem's citywide economic council board of directors.

While managing Jnext, I was recruited as an adjunct lecturer at the Bezalel Academy of Art and Design where **I designed and taught a new semester course on the intersection of technology and architecture**. What began as an elective class morphed into a core curriculum course for all 3rd year architecture students. I was also invited this past year to teach at the competing design school, Shenkar, where I taught 45 students in addition to the class of 75 at Bezalel. Furthermore, I have given guest lectures at the Technion Institute of Technology, Tel Aviv University, and the Hebrew University of Jerusalem, in addition to other local colleges. Teaching has given me the opportunity to share my knowledge within and outside of the classroom, especially with young women who I mentored informally, helping students segway from architecture to STEM adjacent fields, like I did.

As my team progressed in the NASA competition, I became eager to increase my background in space and STEM. This led me to the International Space University's Summer Space program in Ireland, which I was able to attend thanks to the Ilan Ramon Fellowship from the Israeli Space Agency. During the summer I brought my economic development experience and design background to a group setting where I lead the design sub-team of the "astropreneurs" project in addition to writing a chapter on conducting market research. We created a visual handbook and industry overview of the space entrepreneurship landscape, which led to a paper I co-authored that was presented at the International Astronautical Congress (IAC 2017).

**Broader Impacts**

I have been fortunate to have early success as a woman working in highly male-dominated professions. The Venture Capital industry in Israel is 4% women, while the percent of female startup founders is only 10%. This gender imbalance **led me to become a passionate advocate of women and minorities in the technology sector**. I mistakenly thought that I had missed my opportunity to work in STEM because I developed this interest later in life and I wanted to prevent other young women from mistakenly thinking that they did not belong in STEM professions.

Therefore, through Jnext I allocated funds and invested time in supporting diversity of all kinds, but especially gender diversity in the technology sector. I coached the founders of QueenB (queenb.org.il), a non-profit that teaches computer science to 8th grade girls. I also guided them through budget restructuring, prepared them for the grant approval committee, secured government

funding for 50% of the cost of the first year with a 3 year renewal clause, facilitated a partnership with a private donor who funded the other 50% of the first year, and helped them comply with required hiring practices of their instructors. The program grew from 3 pilot schools in Jerusalem to activities across the country. I also funded and served on the advisory board (2015-2016) of SheCodes, a non-profit organization that teaches professional women to code and helps junior women coders improve their skills. The program boasts 20,000 members and 30 active branches.

Because of my activism promoting women in STEM, I was selected to partake in the EU Hypatia Horizon 2020 program, (part of the 'Science, it's a Girl Thing' campaign). Israel was one of the 14 hubs chosen to participate, and I was one of a handful of representatives from Israel sent to a conference in Geneva to develop a new toolkit for STEM curriculum for girls. After this, we organized focus groups with industry and academia to adapt the conference conclusions locally, and with the Bloomfield Science Museum we produced an openly available curriculum source in Hebrew with gender-inclusive science experiments which was adopted nationally.

Even before I had the power and influence to make such impactful contributions as mentioned above, I was passionate about inclusivity in technology. As a college student, I was active as an alumna of the MEET program (Middle East Entrepreneurs of Tomorrow), a highly selective program that brings together Israeli and Palestinian youth to learn computer programming and entrepreneurship skills together. Through speaking at fundraiser auctions, investor outreach workshops, and graduation events I contributed to raising $300K of private donations for the program.

In the past two years, my outreach focus has shifted to space science, as I have witnessed first-hand what an amazing tool space is to get people (and especially kids) excited about science. Specifically, I have mentored educational groups and provided components of science curriculum for youth STEM education. Most recently, I mentored a First Lego League robotics team #1083 who won the best mechanical design award in the international finals. I am also a curriculum contributor to the Young Astronauts program at the Weizmann Institute of Science, the Youth Cadets of Science at the Hebrew University, and PICO Kids, which all provide science educational activities to youth.

At Caltech, I have continued my involvement in community outreach and supporting women in STEM through joining the Caltech's Women Mentoring Women (WMW) club and PCC-Caltech Connection, which brings students from Pasadena Community College (PCC) to Caltech for lab tours, career panels and professional advice. Lastly, I am also a Diversity Ambassador through the Caltech Diversity office to support my peers who are in the LGBTQ+ communities.

**Future goals**

One of the greatest challenges of my generation is combating climate change and create a sustainable future for the next generations. My PhD research at Caltech will further this agenda by advancing the fundamental science of sustainable materials. This work will provide the foundation for inventing new construction materials and producing the future building-blocks of sustainable urban developments. If honored with the NSF fellowship, I will gain flexibility to **pursue ambitious interdisciplinary research in material science and civil engineering to combat global environmental challenges**. It will also enable me to continue supporting women from non-traditional backgrounds going into STEM. With my toolbox from Caltech and the support of the NSF, the bridges of possibility are limitless.

## Intellectual Merit:

**Introduction:** I am interested in developing new sustainable and biodegradable materials that are derived from algae and agricultural waste. The objective is to obtain new materials with properties comparable to commercial plastics and wood but derived entirely from recyclable resources.

While materials such as polymer composites have been thoroughly studied and used widely over the past 30 years[1], development of biocomposites and bio-derived matrices is more recently being investigated[2]. Biocomposites are a type of composite material with one or more of their main components made from sustainable resources. Current biocomposites mainly contain hard to recycle petroleum derived plastic matrix materials. These non-renewable matrix materials have material life cycle disadvantages since the after-life use of most of these biocomposites becomes waste that is incinerated or disposed of in landfills[2]. Fully bio-derived renewable matrix materials could provide increased sustainability. However, these composites still have many issues such as: poor material durability and mechanical properties, costs of production, and materials after-life[2]. A downside of existing fully bio-derived composite (like mycelium) is that their relatively low mechanical properties render them non-sufficient for structural and engineering applications[3].

**Method/preliminary results:**

Our research group developed a new class of biocomposite materials based on cultured, dehydrated and undifferentiated BY-2 tobacco cells, generating preliminary results that support my proposed research[4]. These materials achieved mechanical performance comparable to structural woods and selected commercial polymers by retaining the native plant cell wall composition naturally secreted by growing plant cells. (Fig.1-D). Initial experiments used an undifferentiated herbaceous plant as a model system. Plants demonstrate a notable range of mechanical properties, with their stiffness and bending strength varying over three orders of magnitude[5], providing a promising direction for further investigation. However, significant challenges exist with this technique (Fig.1-C), since the plant cells are slow and expensive to culture, and the dehydrated compression process is not scalable[6].

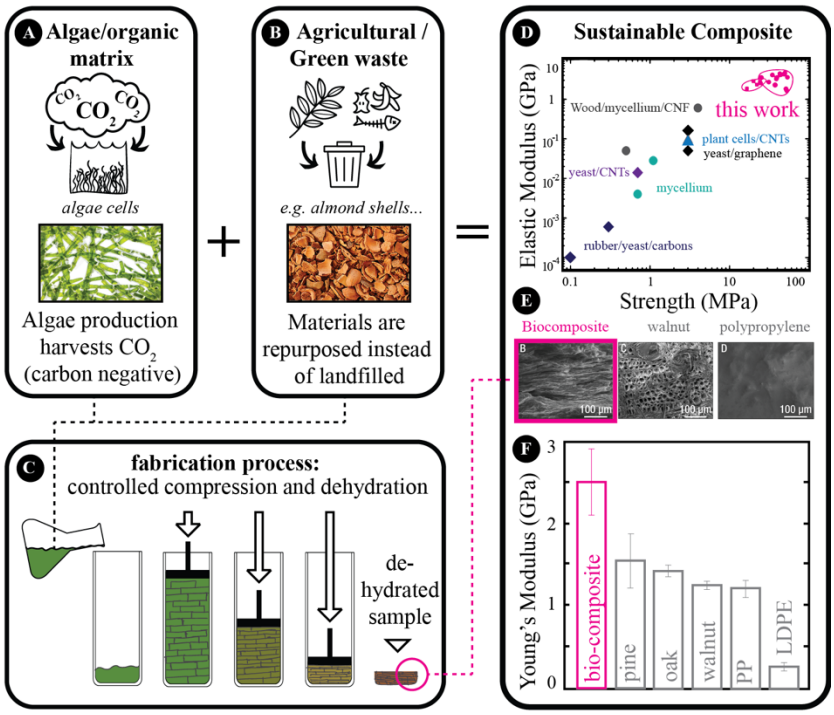



**Fig. 1. Schematic Diagram of the Proposed Approach:** (A) and (B) show the fundamental building blocks/initial materials, (C) is the fabrication process, (D) preliminary results in[4], (E) shows microstructure comparison of biocomposite to reference materials, and (F) shows mechanical properties (Young's Modulus) of this[4] biocomposite is superior to available materials

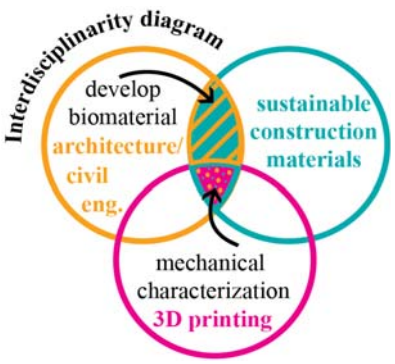
In this proposed research, I will provide a compelling building material alternative by further developing this new class of biocomposites. By leveraging and modifying the group's fabrication techniques, I will use algae instead of plant cells for the bio-matrix (Fig.1-A), in combination with various types of organic waste (Fig.1-B).

**Research plan:** The overarching concept for the new composite realization is to take advantage of the algae and plant cell's capacity to synthesize intricate multi-lamellated structures of cellulose, hemicellulose, lignin and pectin in their cell walls, and thus generate a natural biocomposite with mechanical performance far superior to other biologically derived materials. I will characterize the mechanical properties of the materials fabricated, performing quasi-static and dynamic tests. I will then compare the results to different types of synthetic plastics and wood of similar density.

I propose a fundamental material exploration, which will analyze different types of plant cells and algae with various types of agricultural waste fillers while studying how these changes affect the mechanical properties of the materials produced. Possible organic waste for experimentation are agricultural waste (almond shells, cabbage and coconut), and green waste (yard trimmings, wood, paper and cardboard) which make up a significant portion of materials currently being sent to landfills[7]. Initial testing in our lab indicates that algae may be used effectively as an alternative to plant cells for the natural biopolymer matrix[4]. There are approximately 70,000 known algae species with significant variability in properties warranting further research[8]. Unicellular algae like *Chlamydomonas* are a natural starting point because they lead to conditions of dehydration and growth that are suitable for the proposed plans. Experimentation with different filler additives is expected to result in the introduction of new properties into the composite and support fine-tuning of the mechanical properties of the materials. I will initially use the existing controlled compression and dehydration technique at room temperature, to generate a large-scale parametric study of various matrices and biofiller combinations. Subsequently, I will experiment with the effects of temperature and other parameters variations on the process.

After optimal selection of materials, I plan to study the rheology of the partially hydrated materials. This will be a step towards using them in 3D printing and extrusion based methods. My prior experience in 3D printing will lend both infrastructure and knowledge to my current research.

**Broader impacts:** Developing composite materials from renewable sources will help us sustainably address the global increase in material consumption. The environmental impacts of the new biocomposite material will help reduce the impacts of climate change because algae consumes carbon in its growth process and can reduce $CO_2$ levels[9], while potentially creating a carbon negative material production process. The algae can also be produced locally, minimizing carbon costs of material transportation. Moreover, using a non-wood-based construction alternative should help decrease deforestation.




This research will lay the groundwork for the development of additional mechanically stable fully biodegradable materials. Furthermore, my background in architecture positions me to effectively understand the materials' real-world applications in furniture and construction. In the future, I plan to collaborate with architects to integrate these scalable 3D printing approaches into applications for smart, sustainable buildings.

**[1]** Kinloch, I.A. et al *Science* (2018) **[2]** Mohanty AK, et al *Science* (2018) **[3]** Haneef M, et al *Scientific Reports* (2017) **[4]** Roumeli, E. et al, arXiv:1909.01926 [physics.app-ph] **[5]** L.J. Gibson, J R Soc Interface (2012) **[6]** R. Di Giacomo, et al *Proc. Natl. Acad. Sci. U. S. A.* (2015) **[7]** *US EPA* (2017) **[8]** Guiry, M. *Journal of Phycology* (2012) **[9]** Kumar, K. et al, *Bioresource Technology* (2011)

Scientific advances from the past century are the only reason I have the prospect of a long fulfilling life. My childhood diagnosis of type 1 diabetes instilled in me the gravity of scientific pursuits. From an early age, I saw how the modern scientific tradition of asking piercing questions, disseminating results, and working collaboratively could positively impact my life and countless others. For this reason, I have long pursued an academic career in which I can give back to the scientific community, encourage developments that flow outward into society, and support the next generation of diverse thinkers and innovators. This pursuit has taken many exciting turns, beginning with a transition from an undergraduate degree in biophysics to a three-year research associate position in the Columbia University Department of Biomedical Informatics. At Columbia, I developed novel data science methods for learning from data collected by clinicians and patients, and worked with a multidisciplinary research team to build human-facing computational tools aimed at bringing equity to underserved patient populations. In fact, I initially learned of this team's work while pursuing a hobby of developing statistical methods that I could apply to my own diabetes health data.

As a second-year graduate student in Computing and Mathematical Sciences at Caltech, I have maintained these biomedical collaborations, and am now augmenting them by engaging in a new line of mathematical inquiry into application-neutral frameworks for bridging the fields of machine learning, data assimilation, and dynamical systems. My new role at Caltech has also brought greater opportunities for impacting others, not only through high-impact applications of my current research, but also through local teaching, mentorship, and science communication.

**Intellectual Merit:** Throughout my undergraduate studies, I consistently fixated on the mathematical aspects of my research projects. I recognized my passion for computational research and began taking more courses in applied mathematics. After graduating, I spent three years honing my skills as a Research Associate in Biomedical Informatics. In the evenings, I attended lectures in real analysis and enrolled in courses in dynamical systems and probability theory. At work, I became heavily involved in a Data Mining reading group, where we presented on recent developments in data science, as well as canonical topics—I led discussions on recent conference papers, Shannon's Sampling Theorem, deep learning, and stochastic filtering. At the same time, I was deeply immersed in mathematically founded research projects that helped me develop keen research and professional skills. My research involvement included 1) developing a data assimilation methodology for glucose prediction and inference, 2) evaluating time-series analysis and machine learning methods for detecting adverse drug effects in electronic health record (EHR) data, and 3) collaborating on data science methods for chronic disease self-management in under-served populations. These experiences led me to work with Caltech Prof. Andrew Stuart on broader challenges in dynamical systems and data assimilation. My current research draws inspiration from biomedical modeling challenges, and I work to systematically address them with a novel application-neutral mathematical framework that unites data assimilation and machine learning perspectives. **The sum of my research involvements highlighted in the paragraphs below has yielded 9 journal articles and 4 conference papers across multiple disciplines.**

*Glucose Forecasting:* While at Columbia, I worked with Dr. David Albers to leverage patient data and mechanistic models of glucose-insulin physiology to produce in-the-moment forecasts of blood glucose dynamics. I drove this work through a research pipeline by a) implementing a non-linear stochastic filtering framework (Ensemble and Unscented Kalman Filters), b) pairing the filters with clinically-validated physiologic models of glucose dynamics, c) performing model selection and model averaging, d) developing and comparing offline parameter estimation techniques (optimization and Bayesian techniques, including Markov Chain Monte

Carlo), e) publishing the entire computational pipeline on GitHub, and f) applying the methods to different data settings, including Type 2 Diabetes and Intensive Care Unit monitoring. This work resulted in numerous publications, including a **methodological paper in PLOS Computational Biology**[1] and a **perspectives paper in the Journal of the American Medical Informatics Association (JAMIA)[2], which recently won Best Paper of the Year on Artificial Intelligence in Health** in the 2019 IMIA Informatics Yearbook. I was awarded first-place poster presenter at an annual event hosted by the Columbia University Data Science Institute, where leaders from industry, government, and academia convene to share new findings and common interests. I also presented this work at SIAM Dynamical Systems 2017, the Mathematical Biology Research Training Group at North Carolina State University, the Oberwolfach Research Institute for Mathematics, and the University of Potsdam Institute of Mathematics. This work instilled in me a passion for using statistical methods (e.g. data assimilation) to blend mathematical models with real-world data, and continues to be the inspiration for my new research directions.

*Temporal Dynamics in Electronic Health Records (EHR):* While at Columbia, I also worked with Prof. George Hripcsak to investigate the potential for lagged linear regression methods to identify physiologic drug responses from millions of coarse-grained electronic health records. I observed that our signal detection ability hinged on various pre-processing steps (e.g. temporal transformations, normalizations) and designed experiments to evaluate their importance. This systematic study, which I **published in the Journal of Biomedical Informatics**[3] and presented at multiple conferences, confirmed our initial hypothesis that sequence-time transformations were important, and also revealed surprising co-dependencies between multiple pre-processing choices.

*Data Science for chronic-disease self-management:* Under the supervision of Dr. Albers and Dr. Lena Mamykina, I evaluated machine learning methods for extracting patterns from patient-collected diabetes data, and worked to translate these methods into patient-facing tools that could bridge health literacy and access gaps in underserved populations. One approach used hierarchical clustering, and resulted in a validated visualization tool for clinical nutritionists, which was published in JAMIA[4]. I also helped establish a method for nutritional pattern detection that uses a non-parametric density estimation method from Tabak *et al.*. I worked with students at Columbia to use this method to power a personalized Amazon Alexa health app called T2D2, which became a **National Finalist (1 of 5 teams) in the Amazon Alexa Diabetes Challenge**. A manuscript on this topic is currently under review with the Journal of Biomedical Informatics.

*Graduate Research:* My work in biomedical informatics instilled an immense appreciation for rigorous methodologies that address real-world problems. For this reason, I sought out graduate study at Caltech with Prof. Andrew Stuart, who has pioneered application-neutral mathematical frameworks for data assimilation and Bayesian inversion. I worked with Prof. Stuart to help formalize a paradigm for constraining the state-space of Ensemble Kalman Filters, which was **published in Inverse Problems**[5] and presented as a poster in the Southern California Applied Mathematics Symposium**.** The success of this work sealed my resolve to build general mathematical tools with wide-reaching impact, and prompted me to reflect more broadly on my experiences in biomedical informatics. Through this wider lens, I observed that identifying discrepancies between model and reality is a common challenge across all application areas. This realization has motivated me to investigate ways of blending machine learning and traditional physical modeling and data assimilation techniques to account for unknown forms of model error. I am **currently preparing a manuscript** comparing three competing approaches for their ability to learn the Lorenz 63 system when only given access to perturbed versions of the Lorenz equations

(i.e. a model with complicated errors) and data from the true system. I will present these findings at the Oberwolfach 2020 Data Assimilation Conference and University of Potsdam Institute for Mathematics.

**Broader Impacts:** As researchers, we must use our platforms to shape society into a more inclusive, diverse, curious, and thoughtful place. During my undergraduate years, I paused my biophysics studies to engage in a study abroad program in Chile focusing on public health, traditional medicine, and community empowerment. Through conducting interviews at an indigenous technical nursing high school, I learned that you could achieve better outcomes for patients by creating a more inclusive and equitable environment. This can be facilitated through intercultural education that includes western science and traditional medicine. This perspective later became crucial in my role at Columbia, where I was developing data science methods to help low-income, low-literacy populations improve their health. While I mostly spent time reading papers, coding, and evaluating (ad infinitum), I valued the importance of understanding our target population. As a result of my study abroad experience, I had the linguistic and cultural fluency necessary to lead focus groups among our Dominican patient population at Columbia, which in turn helped me better understand their needs and the goals of the analytical tools I was developing.

To directly promote diversity in research, I am now **collaborating with Mission Unstoppable, a new television series dedicated to inspiring young girls to pursue STEM careers**. I support creative content developers by brainstorming interesting and inspiring segments, drafting explanations of scientific concepts, and editing their scripts to produce a final product that is accurate and engaging. In an effort to further highlight marginalized voices in STEM, I have connected the show's producers to underrepresented scientists at Caltech and Columbia whom I know share similar commitments to diverse representation.

I also aim to engage diverse populations through teaching. While at Columbia, I led a group of seventh-graders at KIPP STAR Harlem College Prep Middle School in a weekly afterschool exploration of computer programming using a curriculum adapted from PlayCodeMonkey.org. I loved supporting students in their self-driven curiosity, and was excited to **use support from the New York Academy of Sciences to bring stimulating STEM lessons to a school with limited resources and an underserved population**. Now at Caltech, I plan to work with Claire Ralph and the Center for Teaching, Learning, and Outreach to support underrepresented STEM students in the surrounding communities of Pasadena.

My roles as a researcher have also brought exciting opportunities for mentoring junior students. As a researcher at Columbia, I **mentored two first-year graduate students** in a project that used machine learning to analyze self-monitoring data and make personalized recommendations regarding nutrition, sleep, and activity patterns—this project resulted in conference presentations and a publicly available Amazon Alexa application. Now at Caltech, I have taken on the challenge of **Head Teaching Assistant for CS 155: Machine Learning & Data Mining**. In this role, I will be responsible for supervising ~20 undergraduate TAs and helping them hone their teaching skills. I will also encourage an inclusive environment where we support students who may be the first in their family to attend college, may be unaware of resources available at Caltech, or may experience hardships that take unexpected tolls on academic performance. I am excited to bring my passion for inclusive, collaborative science and my rich research experience to new communities at Caltech and beyond as I embark on my academic career.

*References:* [1]Albers et al. PLOS Comput Biol. 2017;13. [2]Albers et al. J Am Med Inform Assoc. 2018;25. [3]**Levine** et al. J Biomed Inform. 2018;86. [4]Feller et al. J Am Med Inform Assoc. 2018;25. [5]Albers et al. Inverse Probl. 2019;35.

The next generation of high-fidelity physical modeling techniques will leverage **a new paradigm that blends the power of black-box data-driven learning with the interpretability and transferability of principled mechanistic models.** My goal is to play a significant role in research underpinning this paradigm shift. Consider a prediction task where the goal is to forecast future states of a dynamical system from typically partial and noisy measurements. When substantial physical knowledge about the system's dynamics is available, the canonical approach for solving this problem is to perform data assimilation (e.g. Kalman filtering or its variants), wherein model predictions are balanced with noisy, partial and possibly infrequent observations of the system to reconstruct the model state. When little is known about the system *a priori*, a black-box machine learning approach is often taken, where the data must entirely inform our understanding of the observed system. However, most applied problems live somewhere between these two extremes—we rarely know nothing about a system, so purely data-driven approaches are inherently impeded by their ignorance to hypothetical mechanisms. At the same time, data assimilation techniques typically assume access to a faithful model of reality, and are thus unable to respond to observations beyond the flexibility of their narrowly prescribed physical model class. *In my graduate studies in Computing and Mathematical Sciences at Caltech, I have begun to work with Professor Andrew Stuart to develop a mathematical framework for data assimilation that unites flexible machine learning approaches and physical modeling techniques to systematically address model error and catalyze improvements in application domains that rely on forecasting complex dynamical systems, including specific application in biomedical and geophysical sciences.*

**Aim 1: Build a framework for blending data-driven and physics-based models.** I plan to pursue a unifying theoretical framework that contextualizes existing hybrid forecasting methods and anticipates new approaches. I have identified three distinct literature trends regarding methods that leverage machine learning and physical models to improve predictions of dynamical systems: 1) residual-based approaches, which cast a supervised learning problem of identifying a function (e.g. gaussian process model) that predicts errors of the physical model based on the measurement data[1], 2) approaches that embed solution operators to hypothesized physical systems *within* the structure of a neural network (e.g. reservoir computers)[2], and 3) methods that regularize neural network training objectives such that solutions must balance their fit to observed data with their fit to data simulated from the hypothesized physical model[3]. I am currently preparing a manuscript that studies these three paradigms in a model-error setting simulated by a perturbed version of the chaotic Lorenz 63 system, and compares their predictive performance, interpretability, and computational efficiency. I find that my newly developed mechRNN, a variant of the Pathak *et al.* model[2], significantly outperforms their reservoir computing approach. During my graduate studies, I will study tradeoffs between approaches (1-3) and develop theory to relate them. This work will guide model practitioners who seek predictive performance beyond the capability of physics-only or data-driven-only approaches. Moreover, it will lay experimental and theoretical groundwork for continued progress in leveraging machine learning in computational mathematics.

**Aim 2: Learning augmented data assimilation algorithms.** Many popular data assimilation techniques, like Ensemble Kalman Filtering (EnKF) and 3DVAR, rely on an iterative prediction-correction scheme, where each algorithm is distinguished for how it constructs an update (gain) function that balances between model predictions and observed data. During my graduate studies, I will leverage machine learning and optimization methods to improve data assimilation techniques and produce new, practical algorithms. In 3DVAR, for example, a constant linear gain operator is chosen *a priori*, typically based on physical knowledge of the system.

Recently, I used stochastic gradient descent (SGD) techniques to perform *online identification* of a gain matrix for partially observed Lorenz 63 that outperformed the theoretically derived gain by Stuart et al.[4] While data-driven approaches for performing this optimization have been proposed[5], online SGD offers the possibility for real-time improvements to 3DVAR filtering. I will pursue further opportunities for learning gain functions for non-linear variants of the Kalman filter via optimization, rather than a probabilistic derivation, promoting empirical study and improvement of their probabilistic framings. This aim will contribute to new frontiers by investigating how data assimilation algorithms can be augmented by machine learning and optimization methods to achieve better forecasts and uncertainty.

**Aim 3: Develop a general data assimilation framework that addresses model error.** Recent advances in machine learning provide hope for systematically addressing generalized model error—one of the major outstanding challenges in data assimilation. However, practical successes of modern machine learning methods are insufficiently formalized, making it difficult to predict their behavior. Nevertheless, the completion of Aims 1 and 2 have the potential to shed light on the interconnections between data assimilation, physical modeling, machine learning, and model error. I will study frameworks that embed entire filters within a machine learning model, as well as filters that contain networks of physical and data-driven models. This aim will contribute to the pursuit of generalized data assimilation techniques that use machine learning and physical models to better exploit available data.

**Intellectual Merit:** The proposed work will lead to both challenging research in fundamental computational mathematics, whilst also providing new practical algorithms that blend data assimilation and machine learning to impact applications; the environment at Caltech is rich in such applications and I will build on the symbiotic relations between applications and theory in developing my work. My success in achieving the described aims will be supported by access to advising from Prof. Andrew Stuart, an expert in data assimilation, and Prof. Yisong Yue, an expert in machine learning. Moreover, my studies at Caltech uniquely position me to iteratively develop these frameworks in close contact with high-impact applications, including geophysical problems framed by Prof. Tapio Schneider of the Climate Modelling Alliance and biomedical problems studied by Prof. George Hripcsak at Columbia University and Prof. David Albers at the University of Colorado. I will discuss preliminary work on Aims 1-2 at SIAM Uncertainty Quantification 2020 and the Oberwolfach Research Institute for Mathematics 2020 data assimilation conference, and look forward to further advancing this line of investigation.

**Broader Impacts:** My proposed studies target general mathematics, but also live in a tight feedback loop with high-impact application domains. For example, my previous research in glucose forecasting has inspired my current work on model error; by addressing the proposed aims, there is potential for further impact on glucose prediction through collaborations I maintain with Columbia. I also engage with geophysical applications through the Climate Modelling Alliance at Caltech, where we aim to improve quality and uncertainty of climatological predictions, which are essential for our society's future. I deeply value openness in science: I consistently publish my work on arxiv, make my code freely available on GitHub (e.g. mechRNN), and present frequently about my current work. Pursuing the proposed aims will also better equip me to serve as Head TA for a Caltech machine learning class, mentor undergraduates, and share meaningfully with mathematics students at Pasadena Community College through the Caltech Center for Teaching, Learning, and Outreach.

*References:* [1]Shi et al. ArXiv181108027 Cs. 2018. [2]Pathak et al. Chaos Interdiscip J Nonlinear Sci. 2018;28. [3]Raissi et al. J Comput Phys. 2019;378. [4]Law et al. Discrete Contin Dyn Syst. 2013;34. [5]Zhu et al. Tellus A. 2000;52.

This time three years ago, I stood atop Humboldt Peak as vermilion rays broke over the sky-carving Crestone Massif. This summit in southern Colorado is named for the luminous scientist Alexander von Humboldt whose curiosity, environmentalism, and understanding of Nature across disciplines make him one of my greatest scientific inspirations. Not only are his principles embedded in modern science, Humboldt connected fundamental insights with the betterment of society and the environment. Likewise, I endeavor to deepen our understanding of enzymes and apply their catalytic capabilities to foster a more sustainable global future for all humankind.

Troubled by our impact on the biosphere, I voraciously read climate change and renewable energy literature as a high schooler, leading me to write a forty-eight-page high school thesis on enzymatically produced biofuels. Nature's ultimate green catalysts, enzymes use earth-abundant metals, function in aqueous solvent, and are uniquely malleable to the engineer's hand. My undergraduate research interests were motivated by this thread, ultimately leading to a two-and-a-half-year term at the National Renewable Energy Laboratory (NREL). Now a graduate student in the laboratory of Caltech Prof. Frances Arnold, I am engineering enzymes to sustainably confront challenges in synthetic and environmental chemistry. After my graduate training, I will pursue a career as an academic scientist, inventing biotechnologies to reduce society's footprint while disseminating scientific advances to inform effective and equitable science-grounded policy.

***Intellectual Merit:*** My fascination in the chemistry available to enzymes and their transformative potential for sustainability prompted my undergraduate biochemistry studies at the University of Denver (DU). I commenced research in Prof. Bryan Cowen's laboratory, where I devised enantioselective syntheses of tetrahydroisoquinoline-based natural products with anti-Alzheimers effects. I took advantage of my access to a 500 MHz NMR instrument to explore the bounds of the technique. This foray into organic mechanisms and NMR provided a basis that I use today to characterize enzymatic reactions and also roused my interest in computation as a complement to experiment. Thus, I pursued a minor in computer science at DU and a summer research fellowship in neuroimaging at Washington University in Saint Louis, where I used fast Fourier transform to decipher transient spatiotemporal patterns of neural activity – microstates – from noisy neuroimaging data. Frustrated with the speed of a preexisting script, I taught myself MATLAB and designed a pipeline that improved efficiency by an order of magnitude. I am using this computational background to efficiently process reams of enzyme data in my current research.

As a visiting researcher at UniversitätsSpital Zürich, I collaborated with Prof. Steffen Gay and Merck researchers to test the effects of a raft of lead compounds on inflammatory responses in rheumatic cells. I cultured primary human cells, delicately extracted RNA, and performed a suite of biochemical assays and quantitative PCR to probe their effects. My culminating seminar stimulated much discussion on the mechanism responsible for the significant anti-inflammatory responses I observed. The foundations of cellular and molecular biology that I learned have empowered me to engineer non-model organisms for enzyme expression. Most importantly, I engaged with researchers from varied cultures, experienced how research is conducted abroad, and witnessed how international teams can be motivated to work toward a common challenge.

Upon returning to Denver, I felt compelled to use the research skills I developed to impact the environmental challenges that had long occupied my thoughts. Reviving my interest in advanced biofuels that was wrought during my high school years, I began research under Drs. Michael Himmel and Yannick Bomble at NREL. Life's diversity of metabolic and lignocellulose degrading enzymes are biochemically intriguing and may unlock carbon-neutral biofuel production. With this in mind, I characterized the effect of glycosylation on the most active cellulase known – *Caldicellulosiruptor bescii* CelA. I drove the project by constructing a genetic system to express

the enzyme in several organisms, including its genetically recalcitrant host. I probed the enzyme via proteolysis assays and circular dichroism and demonstrated that it was unequivocally stabilized by glycosylation. My insights inspired a series of molecular dynamics simulations that I oversaw, and I composed results from an 18-member interdisciplinary team into a manuscript. This work represents one of the first investigations of glycosylation in anaerobic bacteria, and resulted in my poster at the *Symposium on Biotechnology for Fuels and Chemicals* and a co-first author article.[1]

I experienced the joy of teaching and collaborating when a PhD student from the Weizmann Institute of Science visited our lab. She was constructing designer cellulosomes, mimics of the multi-enzyme complexes that excel at depolymerizing biomass. My study of CelA glycosylation sparked her to express and glycosylate cellulosome components in *C. bescii*. I trained her in *C. bescii* engineering while she taught me about cellulosome activity. Fusing our areas of expertise enabled us to engineer the first hyperthermostable designer cellulosome, resulting in two articles (one in review)[2,3] and underpinning her PhD thesis. The expertise I gained at NREL resulted in two additional publications,[4,5] my poster presentation at *RosettaCon*, several forthcoming papers, and culminated in my writing a first-author book chapter on multifunctional cellulases.[6] In addition to grounding me in enzyme biochemistry, my time at NREL instilled the importance of teaching and interdisciplinary collaborations in enhancing the impact and reach of science.

I chose to pursue my PhD at the California Institute of Technology for its fearless pursuit of grand challenges and encouragement of interdisciplinarity. Once at Caltech, I pursued an interdisciplinary project I had long been interested in – few structural studies of cellulosomes have been undertaken due to their conformational flexibility, and I saw an opportunity to characterize them via cryo-electron microscopy. I conceived, proposed and initiated this project during my rotation in Prof. Grant Jensen's group. This experience in cryo-EM will synergize with my current research by allowing me to structurally characterize enzymes and their small molecule products.

My work on enzymes inspired an interest in directed evolution, a protein engineering technique that relies on the Darwinian process of mutation and selection to alter protein function. As such, I rotated in the laboratory of Prof. Frances Arnold, who was awarded the 2018 Nobel Prize in Chemistry for pioneering directed evolution. During my rotation, I expanded the ability of tryptophan synthase (TrpB) to form noncanonical amino acids (ncAAs) – molecules that serve as probes for chemical biology, synthons for pharmaceuticals, and might enable bioremediation via biocontainment. However, chemical ncAA syntheses rely on harsh and unsustainable reagents and exhibit limited scope and selectivity. I evolved TrpB to accept nitroalkanes and form novel ncAAs, using my biochemistry expertise to characterize the kinetic rate enhancements of the evolved enzyme over the native enzyme. This study is the first expanding TrpB's scope beyond tryptophan analogues and the findings are published in *ACS Catalysis*.[7] Also during my rotation, I initiated a high-risk project evolving TrpB to form quaternary stereocenters, useful motifs that are difficult to access via chemo- and biocatalysis alike. The engineered TrpB is one of the first biocatalysts capable of forming these motifs enantio- and regioselectively. My resulting manuscript has been posted to *ChemRxiv* and is in revision in the *Journal of the American Chemical Society*.[8]

Reducing our collective impact on Nature is one of the most challenging yet promising frontiers for enzyme engineering. Enzymes are ideal for this challenge as they excel at catalyzing a bevy of defiant transformations due to their selectivity, control, engineerability, and sustainable features. As detailed in my research statement, I am engineering cytochromes P450 to degrade organosiloxanes; the resulting P450s may be used for bioremediation of these (eco)toxic compounds. I will also expand this idea to engineer enzymes for new-to-nature biodegradation of other recalcitrant pollutants. To facilitate discovery of activating mutations and improved enzyme

variants, I am pursuing machine learning methods. Armed with my predoctoral training in research, teaching, and writing, I am eager to push the bounds of this frontier.

***Broader Impacts:*** As scientists, our work takes on its greatest impact when we incorporate myriad viewpoints and peoples into the scientific community. Furthermore, fostering societal understanding of science prompts government to support science and its findings. Realizing this as an undergraduate, I mentored middle school students as a science and mathematics tutor in underserved communities in Denver. Several of my former mentees have now begun college as STEM majors. While in Zürich, I engaged with researchers from dozens of countries and learned how conducting science across cultural boundaries ultimately enables greater impact. These experiences ingrained the importance of a diverse and politically engaged scientific community as well as the imperative to incorporate marginalized populations into scientific education and outreach. Now at Caltech, I am inviting underrepresented faculty to speak here via the Diversity in Chemistry Initiative. I will continue to foster an understanding of science in disadvantaged communities with thoughtful teaching, outreach, and collaborations throughout my career.

My work at NREL and Caltech holds promise for abating our impact on the planet and biosphere. Yet political barriers impede progress toward adoption of renewables, comprehensive environmental action, and scientifically informed government. As such, I have actively learned to effectively communicate science to policymakers. As an undergraduate, I led a group of students in crafting policy briefs on energy and health care with the Roosevelt Institute at DU. At NREL, I led tours of my lab for Senior Department of Energy Officials from the Obama and Trump administrations, conveying our work's impact on science, sustainability, and national competitiveness. Continuing my work at Caltech, I participated in a week-long science policy practicum with seismologist Lucy Jones, where I communicated local water-quality issues to Los Angeles Councilmembers and U.S. Representative Adam Schiff's Chief of Staff. Their enthusiasm to hear from scientists prompted me to travel with three Caltech graduate students to Congress and engage with a bipartisan array of Senate and House offices on energy policy, the future of synthetic biology, and bioremediation. As a result of this trip, I regularly update U.S. Representative Diana DeGette's staff on progress in biotechnology. I will aid progress via my platform as a scientist by continuing to advise policymakers and educate the public via social media and outreach. A diverse and scientifically engaged public both benefits from science-grounded policy and also encourages its development. As such, I will interweave community outreach and science policy to enhance this synergy and realize broad scientific and societal impact.

***Future Goals:*** Upon completing my PhD, I will pursue an academic research career. My future lab will strive to extend our knowledge of enzymes – and bring the catalytic prowess of life to chemically intriguing problems in energy & chemical production, synthetic biology, and biological remediation. At Caltech and beyond, I will augment my science policy experience with policy training and active engagement and harness it to inform policy makers on effectively integrating scientific advances into rigorous policy. Coupling my research and science policy experience, I will contribute novel fundamental insights, knock down technological gates, and advance the nation's security, prosperity, and scientific leadership. Receiving a fellowship from the National Science Foundation will bolster my training, connect me with an ambitious community of scholars, and empower me to contribute to a more sustainable and equitable global future.

**References:** [1] Chung, D[*]; **Sarai, N. S.**[*] *et al.*, *ACS Sustain. Chem. Eng.* **2019**, *7*, 4792–4800. [2] Kahn, A. *et al.*, *Biotechnol. Biofuels* **2019**, *12*, 44. [3] Kahn, A. *et al.*, (In review: *FEBS J.*) [4] Brunecky, R. *et al.*, *Biotechnol. Biofuels* **2018**, *11*, 22. [5] Sammond, D. W. *et al.*, *Biotechnol. Biofuels* **2018**, *11*, 189. [6] **Sarai, N. S.** *et al.*, In *Comprehensive Biotechnology, Third Edition*; **2019**; Vol. 3, pp 14–23. [7] Romney, D. K.; **Sarai, N. S.**; and Arnold, F. H. *ACS Catal.* **2019**, *9*, 8726–8730. [8] Dick, M.; **Sarai, N. S.** *et al.*, *ChemRxiv* **2019** (In revision: *J. Am. Chem. Soc.*).

## Personal, Background, and Future Goals Statement

**Personal background:**

I was raised in Baghdad during the aftermath of the Gulf War in 1990 and associated sanctions with limited resources to explore my interests. Despite this, my passion for the sciences bloomed during my childhood, as I spent my free time reading my father's collection of old encyclopedias and watching the few documentaries available on the national TV. I attended a magnet school, but science communication in Iraq was exclusively theoretical. I therefore advanced through school as a very curious child with many half-answered questions. Due to the devastation of the 2003 war that initiated a cascade of sectarian violence, my family had to flee to Aleppo, Syria, where I finished high school on a scientific track and from that point onward, I have been dedicated to pursuing a **career in science**.

I vividly remember having conversations with my father about his dreams of studying abroad. These conversations inspired me, from a very young age, to think about this possibility for myself, especially as my options after graduating high school in Syria were limited by ongoing instability and lack of resources. When I learned about the Iraqi Students Project, whose goal was to improve English language skills and prepare Iraqi students to apply to colleges in the United States, I moved to Damascus and enrolled in this program. A year of hard work paid off; I was admitted to the entering class at Smith College in Northampton Massachusetts with full financial aid. In the summer of 2012, the situation in Syria had become increasingly dire as in that year, the Syrian revolution was well underway. Sadly, it was impossible for me to see my family before I left for the United States.

As a first-year student in a prestigious liberal arts college, I was encouraged to expand my knowledge base beyond the sciences, and work on my critical thinking skills. While I focused on my studies and acculturating myself to this new environment, my family was escaping yet again from a war zone, in Aleppo. Thankfully they made it to safety by the end of my first year. The transitions were challenging for all of us— difficult yet exciting as I continue to build my life here, away from my family.

At Smith, I also had the opportunity to develop my communication and leadership abilities by working as a teacher's assistant of introductory chemistry labs at Smith in my sophomore and junior years. While my research interests have since evolved through various research opportunities, first with physical chemistry, then incorporating molecular biology and protein biochemistry, the combination of those experiences, in turn, have inspired me to pursue academic research and a potential career as an academic.

**Intellectual Merit**

At Smith College, I majored in Chemistry and minored in Italian Literature. I spent my undergraduate semesters and summers obtaining research experience. I started research in the lab of Professor Katherine Queeney at Smith the summer after my freshman year. The Queeney lab studies the properties of silicon surfaces and how their nanotopography affects the formation of bacterial biofilms, which has important implications to generation of safe medical devices. The lab implements aqueous processes to etch the surface of the silicon to create diverse, functionalized, nanoscale patterns. I etched and silanized surfaces in the vapor phase under nitrogen, and then measured the thickness of the formed layer. I then adhered small biomolecules such as poly-L-lysine and monitored their behaviors on various topographies. I also attempted to optimize the detection of layer thickness based on their respective refractive indices.

This experience introduced me to material science concepts and propelled me to pursue biological applications of my work.

The following year, I applied to the Howard Hughes Medical Institute (HHMI) Exceptional Research Opportunities Program (EXROP). This program allowed me to spend two summers (2014 and 2015) in the University of Utah in the laboratory of Professor Baldomero Olivera. The Olivera lab introduced me to the intersection of chemistry and biology, which I have been passionately pursuing ever since. The lab interrogates signaling in the nervous system using peptides that specifically target ion channels and receptors. The peptides ('conotoxins') are initially derived from the venom of a family of marine predators called cone snails, then further engineered. The first summer, I selectively amplified and sequenced the O1-superfamily of conotoxins from the cDNA of related species. My second summer, I further characterized the conotoxins I had previously cloned. Julita Imperial, a postdoctoral fellow I worked with in the lab had generated peptides using solid phase synthesis, which I then cleaved from resin, optimized folding reactions, and purified them using HPLC. The activity of the cloned peptides was then tested in mice and compared to their native counterparts. I presented my research at the annual meeting for EXROP students at the HHMI headquarters in Maryland in 2015, and my research resulted in publication in the journal *Toxicon*, where **I am a second author.** This opportunity gave me invaluable experience with molecular biology and biochemistry that I continue to benefit from today.

After graduating from Smith College, I took a position as a research assistant in the lab of Professor Jonathan Abraham at Harvard Medical School. The lab is interested in developing antibody-based therapies against viruses that cause human hemorrhagic fevers, including Ebola, Marburg, and arenaviruses. Arenaviruses are rodent-borne agents that cause highly lethal diseases when they are transmitted to humans. My project involved isolating monoclonal antibodies (mAbs) from the blood of survivors of infection by the arenavirus Junín (the cause Argentine hemorrhagic fever). I started by investigating vaccine-elicited antibodies. I continued by subcloning, expressing, and purifying recombinant human monoclonal antibodies, and with these molecules, I carried out functional assays including ELISA binding studies and viral entry neutralization tests (using a pseudotype virus GFP-reporter system and FACS). I finally worked on characterizing antibodies isolated from the blood of survivors of Argentine hemorrhagic fever and establishing a platform to study, in addition to viral entry neutralization, antibody-dependent cellular cytotoxicity (ADCC) activity using a natural killer cell-based assay. **I am an author on a published manuscript** describing the work in *Nature Communications*, as well as **a prospective first co-author** of a subsequent follow-up publication. Thanks to this research opportunity, I am comfortable working with cell cultures, experimental design and protocol optimization that allowed me to be successful in my current project in graduate school.

Currently, I work at the laboratory of professor Lior Pachter, where I am using my varied set of skills to synthesize, optimize and disperse a new method of barcoding hydrogel beads for use in single-cell RNA sequencing, as well as learning bioinformatics; an essential skill in a data-heavy field. This project enables me to delve further into my interest of applicational, interdisciplinary research in a cutting-edge field.

Throughout my academic career, my main goal has been to continuously learn. For example, I have taken classes and participated in research projects not only because they are

within my major, but because I believed they would help me grow a well-rounded set of skills and spark thinking outside the box. Thus, I have become a competitive candidate for interdisciplinary research that values both breadth of knowledge and expertise.

**Broader Impacts**

I am anxious about public speaking; a trait, I have noticed to be more common in non native English speaking women. However, I challenge myself at every opportunity to present my research, whether in laboratory meetings, poster presentations, talks, and conferences. I am motivated to share the hardships and triumphs of my research with other scientists, which provides me with helpful insight, and allows others to potentially benefit from my experience. I am privileged to have **presented my work in a poster with the Caltech community** this past September in the Annual Biology and Bioengineering (BBE) Retreat - a space to connect with other scientists in the division - where I was able to interact with and recruit several first year graduate students, as well as making two collaborations to apply my project and design a panel of custom targeted assays for non-model organisms.

In addition to sharing my research, I volunteer at a wonderful program called Visiting Scientist that Caltech organizes to go once a week and **teach science at a local elementary school**. I started by showing the students my collection of cone snail shells and explained how I studied their "poison", while they were exploring the colors and textures of the different snails. I regularly provide live demonstrations along with the rest of the team, help the students learn key words relevant for the topic, and train them to use a notebook to record their findings. My favorite grades to teach are K-2 as they condition me to distill my knowledge down to the simplest phrasing possible to convey the ideas to them; a skill that has benefited me more than I ever could have anticipated.

Teaching kids is one of my favorite ways to spark scientific interest in young people but the way I have done most impact is in my continuous efforts to recruit and **provide a welcoming environment in the laboratory for women.** I started this at Smith college encouraging my friends to take science classes, apply for fellowships and awards such as EXROP, support students during TA hours, and most importantly encourage women to rotate and join my lab at Harvard, and now at Caltech. I owe a lot of my success to many women who encouraged me to pursue science like my undergraduate advisor Prof. Kate Queeney, and I want to have the same positive impact.

In addition to excellent academic training, at Smith I had the opportunity to develop my **communication and leadership skills** by working as a teacher's assistant of introductory chemistry courses, and I continue to do so at Caltech.

**Conclusion:**

I have faced many obstacles in my pursuit of higher education; I am grateful for these challenges as they have brought me to where I am today. My journey—physical, academic and cultural— has inspired and strengthened me. I traveled across an ocean, leaving my family behind in pursuit of a rigorous undergraduate education. Winning the EXROP fellowship was instrumental; it enabled me to immerse myself in scientific research and propelled me to pursue the next level of higher education. Winning the NSF grant would aid me in my journey to successfully achieve my research goals. I am continuously looking for appropriate and relevant avenues for scientific outreach and collaborations for my project, and I wish to obtain the best training possible for designing scientific technology.

# Graduate Research Plan Statement

Single-cell RNA sequencing (scRNA-seq) has quickly become an essential method of parsing heterogeneous expression profiles of cultured cells, tissue biopsies, and even whole organisms. Several methodologies to separate single cells are currently in use which differ in accuracy and scale; using microfluidics methods to encapsulate individual cells while the RNA is barcoded and reverse transcribed to cDNA was instrumental in increasing the scale of experiments to processing thousands to hundreds of thousands of cells. Consequently, the need for a dependable and reproducible method of labeling single cells prompted the development of several costly commercial kits, such as 1CellBio and 10x Genomics, that provide proprietary reagents and instrumentation to label cells with DNA-barcoded hydrogel beads.

**Maximizing the synthesis efficiency of hydrogel bead barcodes can dramatically increase the scale of scRNA-seq experiments and advance the entire field.** Most commonly, the bead barcodes consist of a distinct series of DNA bases (the "cell barcode"), a unique molecular identifier (UMI), and a poly-deoxythymine (poly-dT) probe to prime the reverse transcription of mRNA transcripts. With the great precision required to label each cell with an individual barcode comes a heavy manufacturing cost to make hydrogel beads with a sufficient variety of combinations of DNA sequences to ensure that each cell is distinguishable in subsequent analysis. Traditionally, building barcodes requires fragments of DNA on the bead to serve as templates for primer extension reactions in several rounds, with washing steps after every segment is synthesized to successfully create a complete single stranded barcode.[1] The required primers and enzymes cost upwards of thousands of dollars, and, due to this cost and complexity, the field has been largely limited to polyadenylated mRNA-capture.

A recent method for synthesizing single stranded DNA (ssDNA) the Primer Exchange Reaction (PER) utilizes a catalytic hairpin loop that templates extension of a target primer followed by dissociation from the product[2]. A binding region hybridizes specifically to its complementary DNA and an exonuclease-deficient DNA polymerase initiates synthesis of the templated DNA. The polymerase terminates synthesis as it reaches a designed stopping point where the polymerase falls off, and the hairpin undergoes branch migration continuously until it detaches, intact, able to repeat the extension again. We have applied PER to synthesize ssDNA on an immobilized surface (bead) for scRNA-seq.

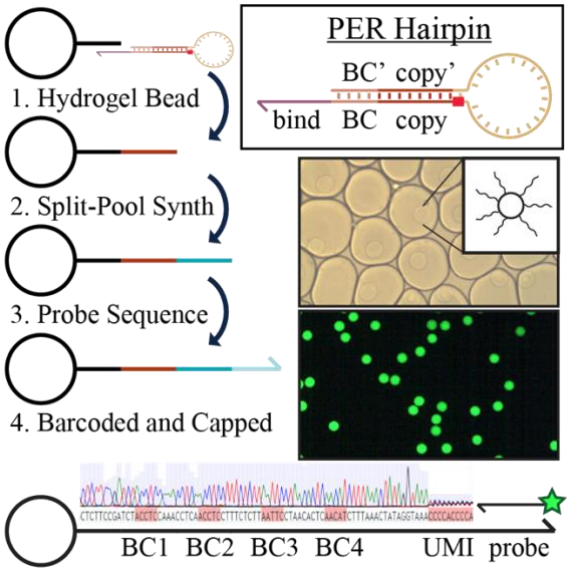


Figure 1: PER barcoding, function, and quality control.

Synthesizing bead barcodes with PER requires 100-fold less DNA compared to previous approaches. Utilizing a split-and-pool method, I increased the number of unique barcodes more than 10-fold compared to existing protocols[3]. I have also confirmed the barcode sequences and probe function by fluorescence and Sanger sequencing, as well as its use in microfluidics-based co-encapsulation with cells for mRNA capture (Figure 1). The objective of my project is to take advantage of inexpensive PER barcoding to enable synthesis of low-cost, customizable hydrogel beads, bringing **open and customizable scRNA-seq techniques within reach of any laboratory**.

**Intellectual Merit**

Thus far, I have developed a method to produce hydrogel beads used for scRNA-seq, leveraging my skillset to combine techniques from synthetic biology and genomics, achieving both an increase in the barcode throughput using less DNA input, as well as an increase in the number of unique barcodes by combinatorial reactions. My degree in chemistry as well as the breadth of my research experience in material science, molecular biology and biochemistry, have given me a great advantage in the experimental design and technology development required to optimize a novel system.

The PER method to synthesize bead barcodes increases the number of molecules produced while significantly decreasing the labor and reagent costs by eliminating the need for intermediate washing steps. This enables users to easily and cost effectively customize the beads by adding any capture sequence to the beads in addition to capturing polyadenylated mRNA.

I will design capturing sequences targeting specific rare genes such as Glycogen Synthase I (GYS1) and Glycogen Phosphorylase L (PYGL) in collaboration with the Thompson Lab at Caltech. **These beads will be used to specifically amplify low-expression genes from peripheral blood mononuclear cell (PBCM) samples that can be lost using standard polyA capture.** The Bjorkman Lab at Caltech is another collaborator for which **I will develop custom capture panels of beads for immune profiling in non-model organisms, an application that is currently not commercially available.** I will also design beads targeting regions in the transcriptome that are usually overlooked such as transcription factors, non-coding RNAs, and exon-exon junctions.

**Broader Impacts**

My goal is to **make scRNA-seq accessible and available to researchers around the world** that do not have the substantial resources which are currently necessary to conduct this research. With this aim in mind, I have, and continue to spread the word about my project by presenting it in conferences, collaborations, and colloquia inside and outside of Caltech. More importantly, I have trained a collaborator from the Eisen lab at University of California, Berkeley on the preparation of PER-barcoded hydrogel beads. My ultimate goal is to establish collaborations to bring open-source scRNA-seq technology to countries in developing areas such as the Middle East, where I can personally connect with scientists and popularize scRNA-seq techniques - where they are largely absent. Making PER hydrogel beads open-source, affordable, and user-friendly has the potential to greatly advance the field, as more scientists would be able to contribute to the potential applications of these beads, which range from expression and immune profiling, to exploratory genomics - especially of non-model organisms - and early cancer diagnostics. For example, my lab is engaged in collaboration with City of Hope hospital to perform scRNA-seq on blood and spinal fluid sample from 21 patients undergoing CAR T-cell therapy.

I have reproducibly demonstrated the quality of the PER method to **make inexpensive, high quality barcoding beads**, and my project's aim to design customizable beads will further contribute to my ultimate goal of **improving the accessibility of single-cell sequencing for the wider scientific community.**

**References**

1. Kishi, J. Y., Schaus, T. E., Gopalkrishnan, N., Xuan, F. & Yin, P. Programmable autonomous synthesis of single-stranded DNA. *Nat. Chem.* **10**, 155–164 (2018)

2. Zilionis, R. *et al.* Single-cell barcoding and sequencing using droplet microfluidics. *Nat. Protoc.* **12**, 44–73 (2017).

3. Zheng, G. X. Y. *et al.* Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* **8**, 1–12 (2017).

During my time in middle school, I read a book, *The Disappearing Spoon* by Sam Kean, that inspired my strong interest in chemistry. After about two years of performing hobby experiments and auditing lectures at Caltech, my parents and I realized that a radical academic acceleration would better help me to achieve my goals. Although auditing lectures was very informative, I could not get hands-on experience in a formal laboratory setting due to liability issues associated with my age. Thankfully, we identified the Early Entrance Program at California State University, Los Angeles as a suitable educational pathway, and I was fortunate enough to make it through the program's highly selective admissions process and begin my studies as a full-time college student at 13 years old. I was soon able to fully enroll in chemistry classes with accompanying teaching labs, and was enthralled by access to modern laboratory equipment. In my third year of undergraduate study, I began conducting research under the supervision of Prof. Matthias Selke, where I focused on singlet oxygen chemistry.

**Intellectual Merit**

Initially, I worked for six months in close collaboration with a graduate student and learned the various methods used in the laboratory's research including air-sensitive techniques, fluorescence spectroscopy, and UV/Vis spectroscopy. I soon transitioned to working independently and even mentored another undergraduate student, studying the photooxidation of triphenylbismuth, a project of my own creation. We performed kinetics experiments to measure the rate at which singlet oxygen, generated by a laser pulse using an organic photosensitizer, was quenched as a function of triphenylbismuth concentration. In 2016, I was accepted as a fellow in the CSULA-Penn State Partnership for Research and Education in Materials (PREM) program, which is dedicated to furthering research opportunities in materials science for underrepresented minorities in STEM. PREM provided valuable support that enabled me to focus on my research. My undergraduate thesis involved studying the photooxidation mechanism of a platinum-bound thiolate complex. Through this project, we sought to gain insight with applications in areas such as small-molecule cancer therapeutics and materials science. My research career began in the Selke lab, and I was thrilled with the opportunity to perform original scientific research and generate knowledge of my own creation as an undergraduate.

Mentoring another undergraduate student was an extremely valuable opportunity: it helped me realize my love for teaching and mentorship that I now hope to pursue in my career. I would like to become a professor and run a research group where I will be able to educate future generations of chemists. My work in the Selke lab gave me practical experience in both inorganic synthesis and physical chemistry, particularly kinetics. However, during the course of my work there, I realized that I enjoyed the synthetic aspects of the research the most, and sought out research experience in pure synthetic chemistry.

As an undergraduate, I spent the summers of 2016 and 2017 performing organic synthesis in the laboratory of Prof. Brian Stoltz at Caltech. My first project, which I undertook as a volunteer in 2016, involved working with a graduate student on drug discovery. In 2017, I was awarded a Summer Undergraduate Research Fellowship (SURF) at Caltech and began working toward a total synthesis of the natural product Melokhanine E. This compound was of interest due to its known antibiotic properties and we believed that a scalable synthetic route would enable further biological studies. Developing a synthesis of Melokhanine E would present a challenge due to the molecule’s complex structure, and required the development of powerful new synthetic methods. Specifically, we envisioned a radical cyclization pathway to access 2-spirocyclo-3-oxindole compounds, and an unusual Bischler–Napieralski-type cyclization to form a new bond at the indole C2 position. I explored multiple strategies for the construction of this compound’s polycyclic framework,

meeting a variety of challenges along the way. In the process, I refined my Schlenk technique and learned to use several powerful new characterization methods, including 2D NMR. Our synthetic strategy had to be revised several times to account for the challenges we experienced during our studies. Eventually, we were able to make considerable progress toward the total synthesis of Melokhanine E, producing a key intermediate only 5 steps away from the natural product. My 2017 SURF culminated in an oral presentation of my research at Caltech's SURF seminar day. The research I conducted as a SURF student cemented my desire to pursue a career in natural product synthesis and synthetic methodology. I was fascinated by the unique logical challenge of building complex molecules from smaller building blocks. In designing a synthesis and carrying it out, I felt like I was assembling a nanoscale puzzle while refining my approach to solving the puzzle in real time.

Last fall, I began a Ph.D. program at Caltech, where I have continued to work under the supervision of Prof. Stoltz. I was also fortunate enough to be awarded a Rose Hills Graduate Fellowship for the 2018-19 academic year. During my first year of graduate school, I opted to work on organic reaction development, with the goal of facilitating the synthesis of useful new compounds. Together with Dr. Alex Sun, I applied the enantioselective decarboxylative asymmetric allylic alkylation methodology developed by our group to the construction of diazepanone heterocycles bearing all-carbon quaternary stereocenters. We were ultimately able to prepare a variety of these heterocycles in high enantiomeric excess and showcase their potential utility in medicinal chemistry through the synthesis of an analogue of the FDA-approved sedative suvorexant. This project culminated in a co-first-author paper, which was recently submitted to a journal. Lately, I have been working on another reaction development project, focused on developing a highly general route to enantioenriched α,α-disubstituted α-amino acids, which have broad applications in peptide therapeutics. I have also been able to study polymer chemistry through a collaboration with the Su group at the University of Arizona, with the aim of developing sensitive devices for chemical warfare agent detection. Furthermore, I have begun to investigate a new total synthesis of a complex marine alkaloid. This exposure to diverse subfields has been highly beneficial to me as a scientist.

**Broader Impacts**

Throughout my undergraduate career, I volunteered at the Huntington Gardens, working on an urban agricultural research project called the Ranch. There, I was responsible for choosing edible plants to grow, caring for these plants and performing garden upkeep, and determining the suitability of our chosen crop varieties for growth on a larger scale in an urban environment. Our aim was to determine which edible plants were most suitable to grow in the urban Los Angeles area, either in backyards or in public gardens, thus serving the local community. My particular interest was in tropical species that presented a challenge to grow in colder climates. This started out as a side-interest, as I have pursued tropical botany as a hobby since grade school, but I came to realize that my work at the Huntington tied in with my prospective career as a synthetic organic chemist. Interestingly, many of the species we aimed to cultivate contain structurally complex and biologically active natural products that present challenges to the synthetic chemist. While at the Huntington, in addition to my garden work, I gave tours of the Ranch to visitors who were interested in applying the information we gathered to their own urban food production. This allowed me to refine my teaching and presentation skills, thus enabling me to better convey scientific concepts to the layperson.

In my graduate studies and in my future career, one of my major priorities is to improve scientific outreach to younger students. From my own experiences in middle school, I know very

well how hard it is to access a real laboratory setting as a minor. For me, the only alternative was to perform experiments at home, which were extremely limited in scope. A marked effort needs to be made to foster the interests of young students who are passionate in science.

With this goal in mind, this fall, I have begun mentoring a group of students from San Marino High School in the Stoltz lab. Together with several of my colleagues, I have been teaching these students some of the foundational concepts of organic chemistry in a small classroom setting. So far, they have developed a good understanding of basic organic functionality and substitution reactions. By regularly completing brief oral presentations, these students both solidify their understanding of the concepts at hand and receive a chance to practice their own science communication skills. These students will soon begin performing original research in the laboratory, where they will synthesize several investigational fragrance compounds. I look forward to continuing to serve as the students' mentor and to the chance to help them refine their laboratory technique and data-analysis skills.

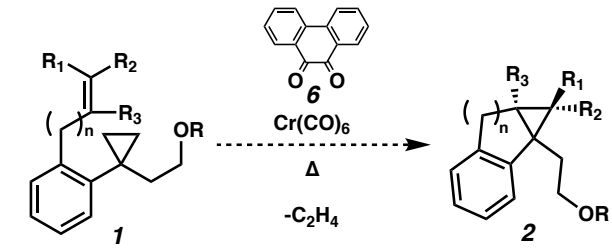
Additionally, in my first year of graduate school, I served as a mentor for an undergraduate student. I was thrilled to be able to expose this student to my research field and to help her refine her laboratory skills, such as her techniques for handling air-sensitive compounds. I also helped her to improve her NMR and mass spectral analysis abilities. Ultimately, she was able to make considerable progress toward the synthesis of a zirconium-based $C_2$-symmetric chiral catalyst for olefin isomerization and polymerization, applications that could directly improve the properties of consumer plastics.

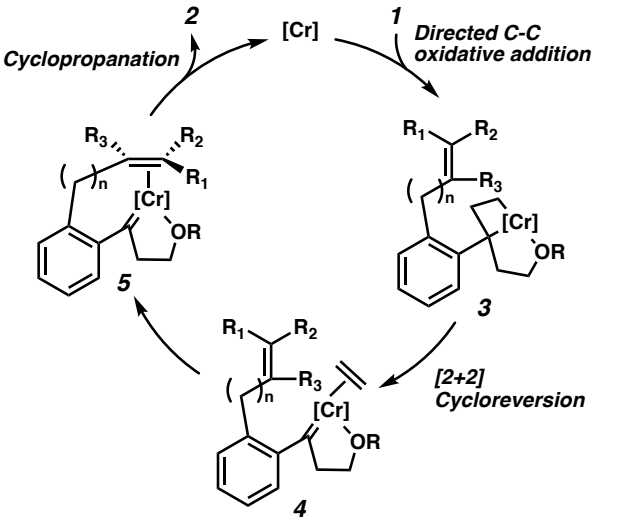
The research I have conducted throughout my undergraduate studies and now in graduate school has far-reaching positive implications for society. The insight into the mechanism of metal-bound thiolate photooxidation we sought in the Selke lab may be applicable to a variety of systems containing metal-thiolate bonds. Platinum-thiolate bonds, in particular, are present in human biology during the metabolism of cisplatin anticancer drugs, and may interact with singlet oxygen generated by cytochrome P450 enzymes. Certain mechanistic pathways for the photooxidation of molecules containing these bonds may lead to the generation of free hydrogen peroxide, which would initiate unforeseen oxidative damage in biological systems. Also, the organic methodology development and synthesis projects I have spearheaded in the Stoltz group will enable more efficient preparation of new small-molecule pharmaceutical substances with minimized side effects and efficacy toward new diseases.

Receiving an NSF Graduate Research Fellowship will allow me the time to focus on achieving my educational outreach goals and to produce research results. During my graduate studies, I will continue to serve as a mentor for high school and undergraduate students, with the hope that I can help ignite the same passion for chemistry that has fueled me to pursue higher education from a young age. After graduate school, my goal is to pursue an academic career in Chemistry, and to eventually run my own research group. A career in academia will allow me to apply my passions for both teaching and research. I have been inspired by my mentors, Prof. Selke and Prof. Stoltz, as well as others whose willingness to extend me the opportunity to participate in research throughout my undergraduate career has so greatly impacted my life. It is important to me to be able to extend the same opportunity to my students in the future. The chemistry program at Caltech provides me with an opportunity to receive a world-class education in total synthesis and synthetic organic methodology. Here, I will be able to adopt the successful strategies of a well-run research group and apply them in my own future research and outreach programs.

**Intellectual Merit.** Metal carbenoids are a class of useful reactive intermediates in organic chemistry, often employed in the preparation of cyclopropane derivatives. These intermediates are commonly prepared from toxic, reactive, and unstable organic diazo species.[1] While several alternative approaches for the preparation of metal carbenoids exist, they typically rely on highly strained precursors that are nontrivial to access, as well as expensive platinum group metals.[2] Versatile methods for the generation of synthetically useful metal carbenoid species that rely on stable, nontoxic, and easily accessible structural motifs as precursors are therefore highly desirable. I propose the development of an intramolecular cyclopropanation reaction utilizing metal carbenoid species generated from simple, readily accessible cyclopropanes;[2] a cyclopropane-olefin metathesis process (Scheme 1). This reaction, catalyzed by earth-abundant chromium, will enable the assembly of complex molecules, including pharmaceuticals, without the use of hazardous diazo reagents or precious metal catalysts.

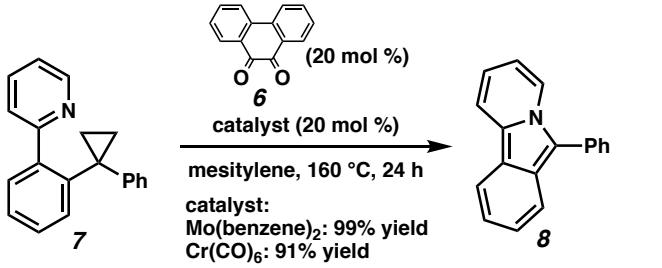
**Scheme 1.** Proposed transformation and catalytic cycle.




Asako et. al. recently demonstrated that zerovalent Cr and Mo complexes, in the presence of ligand **6**, can catalyze the directed retrocyclopropanation of readily accessible pyridyl biaryl substrates, followed by trapping of the generated metal carbenoid (Scheme 2). Although this report demonstrates a promising strategy for metal carbenoid generation, it is only suitable for the synthesis of [2,1-*a*]pyridoisoindoles. In the proposed reaction (Scheme 1), an ether would serve as a directing group, enabled by the high oxophilicity of Cr.[3] The generated metal carbenoid could thus react with an olefin, generating a new, highly substituted cyclopropane. While Mo catalysts gave slightly better yields in Asako's case, the use of Cr carbenoids for cyclopropanation is known to be more efficient than the use of analogous Mo carbenoids.[4]

**Scheme 2.** Retro-cyclopropanation developed by Asako et. al.




Two existing reactions have a degree of similarity to the proposed cyclopropane-olefin metathesis. First, in 1976, Gassman reported the only existing example of a cyclopropane-olefin metathesis, utilizing a W catalyst.[5] However, the reaction was only applicable to simple cyclopropanes and electron poor olefins, and the highest yield obtained was only 14%. Also, precious metal catalysts can generate metallocarbenes from cyclopropenes (which are more strained than cyclopropanes), but these catalysts are expensive, and cyclopropenes are often

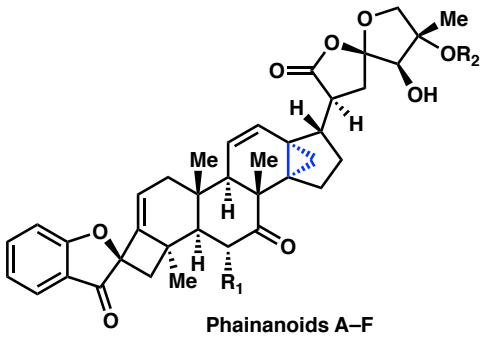
nontrivial to install.[6] The reaction proposed herein would thus represent a considerable advance in the field.

The catalytic cycle for the proposed transformation begins with directed, regioselective oxidative addition of the chromium catalyst into a C-C bond of the cyclopropane to produce metallocyclobutane **3**. Next, a [2+2] cycloreversion, similar to the mechanism of olefin metathesis,[2] will produce key metal carbenoid **4**, stabilized by an adjacent aryl group. Entropically driven release of ethylene gas results in **5**,[2] and a final cyclopropanation will yield the product **2**. Notably, **2** should not react significantly with the metal catalyst: cyclopropanes bearing more substitution are far less reactive in the pyridyl-directed system.[2]

Initial efforts will focus on development of the intramolecular cyclopropane-olefin metathesis shown in Scheme 1. Ethers will be tested initially as directing groups due to the high oxophilicity of Cr, but alcohols, amines, and phosphines are also good candidates. The electronics of the directing group are likely to have a strong effect on the outcome of the reaction. Additionally, an exhaustive ligand screen can be performed – at Caltech, I have access to the Caltech Center for Catalysis and Chemical Synthesis, which contains a high-throughput screening facility. If successful, the reaction can then be applied to a wide array of substrates to evaluate its scope. The experience in synthetic chemistry I have gathered as an undergraduate and now in my time at Caltech has given me the skillset to undertake the development of this new reaction.

**Broader Impacts.** The development of a synthetically useful cyclopropane-olefin metathesis reaction will ease the synthesis of a wide variety of cyclopropane-containing organic molecules, including natural products, pharmaceuticals, agrochemicals, and materials with novel physical properties. For example, the strongly immunosuppressant phainanoid natural products contain a highly substituted central cyclopropane ring, the synthesis of which could be enabled by the proposed methodology.[7] A rapid synthetic route to these molecules and the subsequent development of synthetic derivatives bearing even higher potency could greatly improve quality of life for organ transplant recipients and patients with autoimmune disorders.

**Scheme 3.** The phainanoid natural products.




Additionally, this synthetic method will allow the use of hazardous diazo compounds to be circumvented in many cases, improving safety and efficiency in the laboratory, and thus preventing numerous tragic incidents. Readily available alternatives to diazo reagents will also improve the feasibility of carbenoid usage on the industrial scale. Furthermore, the relative abundance and low price of chromium compared to noble metals will increase sustainability in organic chemistry.

**References.**


1. Jia, M.; Ma, S. *Angew. Chem. Int. Ed.* **2016,** *55,* 9134–9166.
2. Asako, S.; Kobashi, T.; Takai, K. *J. Am. Chem. Soc.* **2018,** *140,* 15425–15429.
3. Kepp, K. P. *Inorg. Chem.* **2016,** *55,* 9461–9470.
4. Miki, K.; Nishino, F.; Ohe, K.; Uemura, S. *J. Am. Chem. Soc.* **2002,** *124,* 5260–5261.
5. Gassman, P. G.; Johnson, T. H. *J. Am. Chem. Soc.* **1976,** *98,* 6058–6059.
6. Archambeau, A.; Miege, F.; Meyer, C.; Cossy, J. *Acc. Chem. Res.* **2015,** *48,* 1021–1031.
7. Fan, Y.; Zhang, H.; Zhou, Y.; Liu, H.; Tang, W.; Zhou, B.; Zuo, J.; Yue, J. *J. Am. Chem. Soc.* **2015,** *137,* 138–141.