BAR-ILAN UNIVERSITY – YESHIVA UNIVERSITY

Summer Science Research Internship Program 2025



The Bar-Ilan University–Yeshiva University Summer Science Research Internship Program is an amazing research opportunity for undergraduate men and women, allowing them to contribute to the forefront of science research taking place in Israel. Generously supported by former chairman of Bar-Ilan's Global Board of Trustees, Dr. Mordecai D. Katz z"l and his wife Dr. Monique Katz, the Irving I. Stone Foundation and the Zoltan Erenyi Fund, students gain invaluable laboratory skills, along with an unforgettable summer experience.

Program Directors: Prof. Arlene Wilson-Gordon and Prof. Ari Zivotofsky
Av and Em Bayit: Rabbi Feivel and Tamar Segelov

TABLE OF CONTENTS

Brain Sciences **4**

Work-Related PTSD Symptoms Following the October 7th Events in Israel: A Pilot Study Comparing Therapists and Media Professionals 4
Adina Feldman

A Taste of Socio-Psycholinguistics 5
Ruti Frohlich and Ilana Pollak

Automating a Pipeline for Processing EEG Data in MATLAB 6
Vered Gottlieb

Engineering **8**

Near-Optimal Byzantine Reliable Broadcast with a Message Adversary 8
Naomi Beck

Thermal-Induced Delay Variation in FPGAs Measured Using Ring Oscillators 9
Shoshi Cantor

Feature Extraction and Model Replication for Hardware-Efficient Deep Learning 12
Ezra Cohen

Multi-Agent Constraints for K-Robust CBS 14
Gabriel Dershowitz

Distributed Computing Simulator 16
Tani Diament

Gene Interaction Analysis Using Abstract Boolean Algebra 18
Andrew Haller and Moshe Wieder

Optically Identifying Bacteria 21
Aviva Klahr

Detection of Atto 532 Fluorescent Dye Using a High Throughput Optical Modulation Biosensing System 24
Rachel Sharon

Testing Natural Oscillator Synchronization on an FPGA 26
Anna Weisman

Life Sciences **30**

A Study of SIRT6 and the Molecular Biology of Aging 30
Ben Epstein and Leah Weiss
Tracking the Stress Granule Dynamics of the RNA-Associated Protein MOV10 33
Ahuva Halpert
Applying Metaproteomics to Investigate the Interaction Between Cancer Tumors and the Gut Microbiome 36
Talia Hazan
Targeted Gene Correction of the Artemis (DCLRE1C) c.1543G>T Mutation in CD34 Cells Using CRISPR-Cas9 and HDR 39
Anna Laufer
Cython Optimization When Creating Gene-Coexpression Matrices 41
Ariel Melnitsky
Modeling CAKUT Pathogenesis Using Kidney Organoids with Heterozygous DACT1 Mutations 42
Rachel Schwartz

Physics, Chemistry, and Mathematics **44**

How Free are the Oceans and Marine Clouds from Continental Influence? 44
Noah Bodner
Radiative Diffusion Utilitizing a Monte Carlo Simulation 46
Nissim Farhy
Over and Over Again: Exploring Iterations of Voronoi Diagrams 48
Hannah Goykadosh
Performance Evaluation of Platinum Catalysts in PEM Hydrogen Fuel Cells 51
Batyah Jasper
Implementing Secure Multiparty Protocols in Python 52
Elza Koslowe
Phosphorylation on the Molecular Level in E2F8 Using NMR Spectroscopy 53
Devora Weinstein

Editor: Elza Koslowe
Contributing Editor: Shoshi Cantor

BRAIN SCIENCES



Adina Feldman, Ruti Frohlich, Ilana Pollak, Vered Gottlieb

Work-Related PTSD Symptoms Following the October 7th Events in Israel: A Pilot Study Comparing Therapists and Media Professionals

Adina Feldman

Advised under Professor Ilanit Hasson-Ohayon and MA student Eliya Galina

Exposure to trauma often extends beyond the individuals who experience it directly, affecting those who interact with survivors — a process referred to as secondary traumatization. This phenomenon is characterized by symptoms similar to post-traumatic stress, including intrusive thoughts, avoidance behaviors, negative alterations in mood and cognition, and hyperarousal, which can emerge in individuals indirectly exposed to trauma. Research has extensively documented secondary traumatization among first responders and healthcare professionals, such as police officers, nurses, and psychotherapists, who regularly encounter survivors of violence and disaster. Despite this, media workers — who frequently engage with highly traumatic material through reporting, photography, and interviews — remain an understudied population, even though their work can place them at significant risk for post-traumatic distress.

The present study investigates secondary traumatization among mental health

therapists and media professionals in Israel in the aftermath of the October 7, 2023 attacks and the ensuing conflict. These events involved large-scale violence, mass casualties, abductions, and sexual assaults, creating an extreme context of collective trauma. Professionals were exposed both personally, through proximity or personal connections, and vicariously, through interactions with survivors, highlighting a "shared trauma" environment. Using thematic analysis of participants' experiences, this study explores how empathy, professional experience, and trauma exposure relate to post-traumatic symptom severity and whether these associations differ between therapists and media workers.

Preliminary findings reveal elevated secondary PTSD symptoms in both groups. Distinct dimensions of empathy predict symptom severity in nuanced ways, and patterns of distress differ between therapists and journalists, reflecting the unique demands and coping mechanisms of each profession. These results underscore the need for tailored interventions and support strategies, offering critical insights into the risk and protective factors that shape secondary traumatization among professionals exposed to collective trauma.

A Taste of Socio-Psycholinguistics

Ruti Frohlich and Ilana Pollak
Advised under Prof. Sharon Armon-Lotem and Prof. Carmit Altman

Socio-psycholinguistics explores the interplay of society, psychology, and language, a framework outlined by Walters' [1] Bilingualism SPPL Interface. During the BIU-YU summer program, we engaged with three primary research studies that together highlighted how children's linguistic development is shaped both socially and cognitively.

First, building on the work of Sari Zarif (PhD student), we studied the effects of narrative intervention among children from low and high socioeconomic backgrounds. In a longitudinal design, participants completed narrative and vocabulary assessments before and after cycles of narrative intervention. Results demonstrated measurable improvement, particularly for children from low SES backgrounds, suggesting that structured narrative support can mitigate linguistic disparities (Figure 1).

Second, under the guidance of Ziva Abate (MA student), we examined the relationship between identity and linguistic proficiency. Hebrew monolinguals, Russian-Hebrew bilinguals and Amharic background speaking children were assessed via standardized measures of phonological awareness, vocabulary, and syntax, alongside questionnaires probing ethnic and societal identity. The findings revealed a clear correlation between Hebrew proficiency and perceived Israeli identity. Specifically, vocabulary and syntax — but not phonology — predicted stronger Israeli identity, underscoring the socio-cultural

weight of language choice beyond its mechanical features (Figure 2).

Finally, working directly with Professors Armon-Lotem and Altman, we contributed to the assessment of results from the MULTI app, designed to efficiently assess developmental language disorders (DLD) in monolingual, bilingual, and multilingual children. By encoding and analyzing children's responses in nonword and sentence repetition tasks, we observed the underlying psychological processes that structure bilingual language use. This tool demonstrates how psycholinguistic methodologies can streamline clinical assessment while accounting for linguistic diversity (Figure 3).

Taken together, these projects reveal the necessity of integrating social and psychological dimensions in the study of language. SES and identity emphasize how external contexts shape linguistic growth, while analysis via tools like the MULTI app reveals the inner processes guiding that growth. Our summer work illustrates that bilingualism cannot be understood in isolation but must be approached through the combined lenses of society, psychology, and linguistics.

1. J. Walters, Pragmatics & Cognition, **8**, 1–26 (2000).

2. S. Armon-Lotem et al., Linguistic Approaches to Bilingualism, **7,** 311–345 (2017).



Figure 1. Example narrative used in an intervention

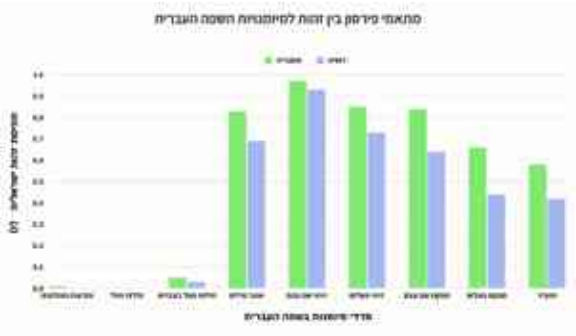


Figure 2. Correlation between Hebrew proficiency and Israeli identity perception

Target item	OM-Omission	SUB -Substitution	ADD-Addition	TRANS-transpositions
magal		nagal		lagam
kriper	kiper;kipe	kiper		
dfula	fula	nefula	defula	
Psycho-encoding				
magal		SUB		TRANS
kriper	OM,OM2	SUB		
dfula	OM	SUB	ADD	

Figure 3. Sample task from the MULTI app for assessing developmental language disorders

Automating a Pipeline for Processing EEG Data in MATLAB

Vered Gottlieb
Advised under Dr. Leehe Peled-Avron

Properly processing and analyzing electroencephalography (EEG) data is a crucial step in establishing a link between

human behavior and corresponding neural processes. Dr. Peled-Avron developed a pipeline that cleans, processes, and analyzes the EEG data collected as part of her research on the neural processes that are triggered by human touch [1]. The process requires the user to create the pipeline for every dataset they wish to analyze, through interacting with UCSD's EEGLAB and UC Davis's ERPLAB, MATLAB toolboxes created for this type of analysis. When repeating this process for data from hundreds of subjects, it not only becomes extremely time-consuming, but also leaves much room for issues caused by human error within the data processing. The objective of this research was to create a MATLAB script to automate the existing EEG analysis pipeline. To accomplish this, I started by familiarizing myself with the existing pipeline, noting updates made within EEGLAB and ERPLAB, along with resolving some issues caused by toolbox development errors. Then, I went through the entire pipeline, using EEGLAB and ERPLAB manually to generate the list of MATLAB commands and functions each tool creates and calls. Then, I fine-tuned those commands and added control flow to create an automated MATLAB script calling each of those functions in turn, while allowing the users to interact with the GUIs to add their specific measurements. Then I integrated additional steps from other pipelines [2] to provide enhanced results. In writing the scripts, I tried several different approaches to balance ease of use, script runtime efficiency, and flexibility. Initially, all the processing parameters were hard-coded. While this made the script fast and user-friendly, it limited the potential uses to below an acceptable threshold. Therefore, my next approach was to enhance the automated pipeline to allow the user to choose configuration files that would direct the data analysis. This provided better flexibility without compromising ease of use. However, it was still too limited for the script's intended use cases. Therefore, I redesigned the script to have the user input the precise measurements and parameters they want in each step through interacting with EEGLAB's and ERPLAB's GUIs. This created a more flexible script, without severely impacting the ease of use. Together, these final changes yielded a pipeline that was much more efficient than a fully manual pipeline and much more flexible than a hard-coded script. The script works as an automated pipeline to analyze the EEG data, allowing users to input not just the data they are analyzing, but the parameters of the measurements taken within each step of the process. This work has greatly contributed to the data processing and analysis within the lab, making it easier, faster, and more efficient.

1. Leehe Peled-Avron and Simone G. Shamay-Tsoory. "Don't touch me! autistic traits modulate early and late ERP components during visual perception of social touch." Autism Research, **10,** 6 (2017): 1141-1154.

2. Cheng Xiaoqin et al. "A pleasure that lasts: Convergent neural processes underpin comfort with prolonged gentle stroking." Cortex, **188** (2025): 13-24.

ENGINEERING



Tani Diament, Andrew Haller, Moshe Wieder, Gabriel Dershowitz, Ezra Cohen, Shoshi Cantor, Naomi Beck, Aviva Klahr, Rachel Sharon, Anna Weisman

Near-Optimal Byzantine Reliable Broadcast with a Message Adversary
Naomi Beck
Advised under Prof. Ran Gelles

Background

In distributed systems, a group of computers (or *nodes*) often need to agree on the same piece of information, even if some nodes are faulty or malicious. Byzantine Reliable Broadcast (BRB) is a fundamental building block that guarantees all correct nodes eventually deliver the same message, or none do. This reliability is critical in many real-world applications such as blockchains, replicated databases, and fault-tolerant cloud systems.

Traditional BRB protocols assume only faulty nodes can misbehave. However, real networks also suffer from message loss: some messages may simply never reach their destination. To address this, Albouy, Frey, Raynal, and Taïani [1] introduced the AFRT algorithm, which extends BRB to handle a *message adversary* — an entity that can deliberately drop some messages in addition to the presence of Byzantine (malicious) nodes. Their work showed that reliable broadcast is possible as long as the number of faults and message losses stays below certain limits.

Building on this, Albouy et al. [2] proposed a new version called the Message-Adversary Byzantine Reliable Broadcast (MBRB)

protocol. This version reduces communication cost by splitting the message into smaller pieces (erasure coding) and using cryptographic tools such as threshold signatures and vector commitments to ensure security and efficiency.

Problem and Approach

While the theory behind the final version of the Message-Adversary Byzantine Reliable Broadcast (MBRB) protocol [2] is well established, its practical implementation is still in early stages. The only available code was based on an earlier version of the protocol, posted as an arXiv preprint [1]. Version 1 used erasure coding and Merkle trees to reduce communication overhead, but it had not yet integrated the newer cryptographic tools introduced in the final version.

My project focused on bridging this gap between the arXiv Version 1 codebase and the conference Version 2 specification. I studied the AFRT algorithm [1] and both versions of MBRB to understand their cryptographic foundations, then began exploring how to extend the existing Python simulation framework. This required identifying how the current code handles message fragmentation and verification (through erasure coding and Merkle proofs) and determining where to incorporate the Version 2 primitives of threshold signatures and vector commitments. I also explored available cryptographic libraries that could support these features in practice.

Conclusion

This project represents an early step toward translating the theoretical efficiency of the MBRB Version 2 protocol into practice. By studying the differences between the arXiv Version 1 code and the final conference specification, and by mapping out an implementation path for vector commitments and threshold signatures, this work lays the foundation for future experiments. The next stages include completing the integration of these primitives into the code, running simulations under different adversarial settings, and comparing empirical communication costs with the theoretical predictions of the Coded-MBRB algorithm.

1. T. Albouy, D. Frey, M. Raynal, and F. Taïani, Theor. Comput. Sci., **978**, 114110 (2023).

2. T. Albouy et al., in Proc. OPODIS (LIPIcs), **292**, 14:1–14:20 (2024).

Thermal-Induced Delay Variation in FPGAs Measured Using Ring Oscillators

Shoshi Cantor
Advised under Dr. Yoav Weizman and Bar-Ilan graduate Avishai Kolet

The global integrated circuit (IC) market revenue in 2025 is estimated at over US$583bn [1] and continues to grow as these devices become increasingly embedded in nearly every aspect of modern life — from smartphones and laptops to automobiles and household appliances. As our reliance on ICs deepens, the demand for reliable and efficient digital hardware

becomes more pressing. Even small disruptions in signal quality can have major consequences. One critical factor affecting the performance and longevity of ICs is temperature. Heat not only accelerates defect mechanisms and long-term device degradation but also directly influences propagation delay, the fundamental time required for a signal to pass through a logic gate. These delay variations can accumulate across complex circuits, undermining timing margins and potentially leading to data errors, instability, or outright system failure. Field-programmable gate arrays (FPGAs), a type of integrated circuit that can be configured (and reconfigured) by a user after manufacturing to perform specific digital logic functions, provide a flexible hardware platform to implement and study these effects in practice.

For the reasons explained above, accurate temperature monitoring and delay characterization have become essential for ensuring reliability in advanced digital systems. A variety of techniques exist for sensing and quantifying thermal effects, ranging from analog temperature sensors to on-chip monitoring circuits. Among these, ring oscillators (ROs) have emerged as a particularly effective and lightweight tool. By stringing together an odd number of inverters in a loop, ROs produce a free-running oscillation whose frequency is inversely proportional to the propagation delay of the gates. The formula is as such:

$$f = \frac{1}{T} = \frac{1}{N*2*\tau},$$

where N is the number of inverter elements within the loop, and τ is the inherent delay of each individual inverter logic element. When temperature rises, gate delays increase, and the oscillation slows. Thus, tracking the frequency of an RO provides a direct, real-time measure of how thermal stress impacts circuit timing.

Over the course of this summer, I was tasked with assisting on designing and implementing a laboratory exercise that will be used within an Advanced Circuit Analysis course in the Bar-Ilan University Electrical Engineering department. The exercise focuses on testing signal degradation and thermal-induced delay variation using FPGA-based ring oscillator circuits. The goal was to create a system that would allow students to directly observe how temperature affects the propagation delay of digital logic, thereby linking classroom theory about reliability and signal integrity to practical, hands-on measurement. I started by learning Verilog (a popular hardware description language) to design a system of ring oscillators and counters that could then be used to program an FPGA board. Once designed I tested the output of my oscillator counter system using GTKWave, a cross-platform waveform viewer for visualizing simulation results from hardware description languages. This gave me an initial understanding of the necessary components of the system as well as how oscillation frequency could be captured through counter-based

measurement, and how noise or glitches might appear in a raw digital waveform. One thing I noticed and accounted for in my code was that if the clock frequency that controlled the counter was higher than that of the oscillator itself a high edge might be double counted. To account for this, I added edge detection code that could hold in memory the previous output of the oscillator during the last rising edge of the counter and compare it to the new value to check if it truly was a new oscillation.



Figure 1. GTKwave output

(w[2:0] = scalar array declared as 3 inverters, clk = clock of counter, din = output of 3rd inverter, din_prev = output of 3rd inverter at last clk rising edge, rst = reset, count[31:0] = 32-bit counter)

After testing confirmed I had a functioning design, to lessen the steep learning curve and cut some of the time necessary to properly learn how to use Vivado (an integrated design environment and software suite by AMD for designing and implementing hardware Systems on Chips & FPGAs), I was provided with a working configuration (designed by Avishai, the student I assisted in the lab) already downloaded on a PYNQ-Z2 FPGA board. Avishai had previously written python scripts that automated the oven heating and temperature controls as well as the overall experiment which called Jupyter Notebook scripts to control the board, sample the counters, and generate data files. With the FPGA and data collection pipeline in place, the focus shifted toward integrating and stabilizing the thermal environment. The FPGA board was placed inside a hot-air oven controlled by a proportional-integral-derivative (PID) feedback loop. Much of the following weeks were dedicated to debugging and refining this system. The oven's warm-up time had been excessively long, the initial overshoot was too high, and the stability window was not tight enough to hold temperature within the desired bounds.

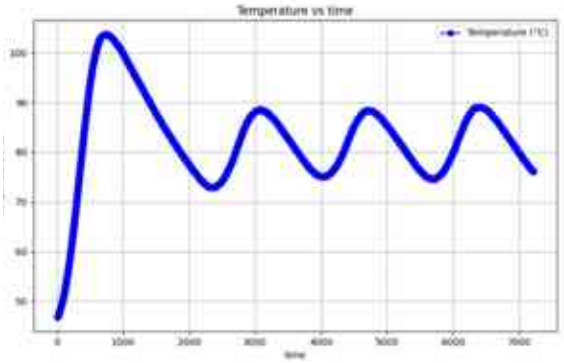


Figure 2. Initial temperature vs. time graph prior to PID adjustments

Iterative adjustments (as well as some amount of trial and error) to the PID parameters were required to achieve steady-state operation suitable for reliable testing across a thermal range of 80 - 120 °C.

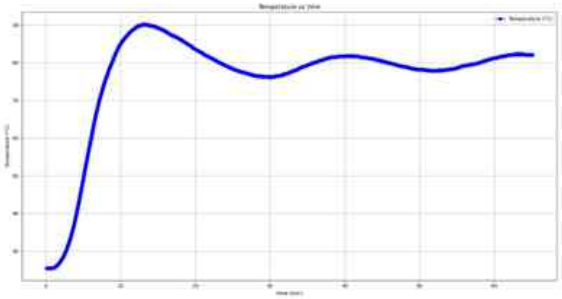


Figure 3. Temperature vs. time graph post PID adjustments

By the end of the project, the combined system of oscillators, counters, data collection scripts, and thermal control had been fully integrated and operational as a lab exercise. The hot-air oven now reaches its target temperature in approximately 10–12 minutes, with an additional 8–10 minutes to settle into steady-state, maintaining a tight tolerance of ±2 °C. Sampling data from all three ring oscillator configurations — NOT, NAND, and NOR — is automatically recorded and stored locally, and system parameters such as number of samples or duration of test can be easily adjusted within the process_main script.

While the program timeframe did not allow for comprehensive measurement of signal degradation across various high temperatures to quantitatively characterize the signal degradation and component delay within the specified temperature range, nor optimize the parameters for such experimentation, the system now provides a flexible and functional platform for such experiments. Additionally, Avishai and I wrote a student manual which explains all aspects of the system to be used within the laboratory course, allowing students to easily use and build upon the system currently in place. By translating the problem of thermal reliability into a hands-on laboratory exercise, this work delivers both a research tool and a practical teaching resource. Students in the Advanced Circuit Analysis course can now directly explore the relationship between temperature, propagation delay, and signal integrity.

1. Statista, Market Insights – Integrated Circuits – Worldwide (2025)

Feature Extraction and Model Replication for Hardware-Efficient Deep Learning

Ezra Cohen
Advised under Dr. Leonid Yavits and MSc student Yuval Harary

This research project explores the internal mechanisms and representational power of deep convolutional neural networks by implementing, analyzing and reverse-engineering residual architectures within the PyTorch framework. Deep learning, a subfield of machine learning, uses multilayered artificial neural networks to automatically extract meaningful patterns from complex, high-dimensional datasets — enabling applications in image recognition, natural language processing, autonomous driving, and beyond. The project began with an in-depth theoretical study of neural networks, focusing on forward and backward propagation, activation functions, and gradient-based optimization. This foundation supported my transition to using PyTorch, a flexible and widely adopted deep learning library.

As a hands-on component, I implemented a custom version of ResNet-8, a simplified residual network architecture. ResNets introduce identity-based shortcut connections that allow gradients to flow

more easily through deep models, mitigating the vanishing gradient problem and enabling training of deeper architectures. After training the model on a labeled image dataset, I focused on extracting features from internal layers. Using forward hook functions in PyTorch, I intercepted the output of specific layers during inference, capturing intermediate feature maps. These features, learned abstract representations of the input, were saved as NumPy arrays. This method, known as feature extraction, underpins transfer learning, which allows pre-trained models to generalize to new problems with reduced training cost and data requirements.

More specifically, a key contribution of this work is its role in supporting efficiency-oriented hardware deployment. While most modern neural networks are trained and executed in FP32 precision, specialized accelerators aim to achieve faster and more energy-efficient inference by using lower-precision formats such as FP8 or FP4. Transitioning to these formats, however, requires rigorous validation to ensure that accuracy is preserved despite reduced numerical precision. The methodology developed in this project, implementing custom networks and extracting intermediate features, provides a reference framework for this validation. By capturing and saving feature maps from a standardized FP32 model, we establish ground-truth outputs against which hardware-executed models in FP8/FP4 can be compared layer by layer. This process ensures that discrepancies caused by quantization or precision loss can be identified and corrected, thereby bridging the gap between high-level software models and low-precision hardware implementations. In this way, feature extraction not only demonstrates the interpretability of deep networks but also serves as an essential tool for enabling efficient, hardware-optimized inference.

In addition to working with ResNet-8, I undertook a model equivalence analysis by mapping PyTorch's built-in torchvision.models.resent18 architecture onto a ResNet-18 that I implemented from scratch. This required a layer-by-layer comparison of architecture structure, parameter counts, forward pass outputs, and intermediate feature maps. To ensure identical behavior, I initialized both networks with the same weights and inputs, then verified that their outputs and internal activations matched. This process provided insight into model internals, debugging strategies, and the fidelity of user-defined architectures. Such mapping is essential for building trustworthy models, modifying architectures for novel use cases, or porting pre-trained weights from one framework or architecture variant to another.

This step is critical for a larger research project of the lab focused on developing a novel hardware accelerator for neural network inference. Since the specialized chip does not natively support high-level frameworks such as PyTorch, neural networks must be manually implemented

and their operations explicitly mapped to hardware primitives. The skills gained from replicating and validating ResNet architectures directly inform this process: they enable accurate translation of model components beyond simple matrix multiplications, including convolution, batch normalization, and residual connections, into hardware-executable operations. In practice, the methodology developed here provides a blueprint for porting complex models to hardware, ensuring correctness while exploiting architectural optimizations that accelerate inference.

Altogether, this project illustrates the practical utility of deep learning models beyond simple end-to-end training: as modular systems whose layers encode powerful and reusable features. It demonstrates how feature extraction and architecture replication support tasks like efficiency-oriented hardware deployment and development of a novel hardware accelerator for neural network inference.

Multi-Agent Constraints for K-Robust CBS

Gabriel Dershowitz
Advised under Dr. Dor Atzmon

Introduction/Background

The Multi-Agent Path Finding (MAPF) problem is defined by a graph of n agents, each of which has a start and goal vertex on the graph. At each discretized time step an agent can either wait at its current position or move to an adjacent location. The solution is defined by a set of paths for each of the agents such that each agent reaches its goal vertex without creating a conflict with another agent.

Conflict-Based Search (CBS) is an optimal algorithm used to solve the MAPF problem. Paths are found for each agent independently using a "low-level" solver such as A*. This solution set is treated as the root node of the search tree. From that root node a conflict is detected in the solution set and the tree is split into two nodes; each side of the tree adds a constraint to the A* solutions that agents x and y, the agents involved in the conflict, cannot be present at position p at time t (the conflict). This process is continuously repeated until a set of paths is found in which there is no conflict amongst the agents.

Disjoint-CBS improves upon CBS. One issue with basic CBS is that the same solutions can be present in multiple nodes. Disjoint-CBS gets rid of this repetition by creating a positive constraint, meaning that the agent has to be at position p at time t, as well as a negative constraint in the other node. This method has been shown [1] to significantly improve the runtime of CBS, as it reduces the amount of calculations the computer must run.

In practice, however, agents/robots may experience unexpected delays (of up to k steps). K-Robust MAPF is an algorithm which adapts CBS to solve this problem. The K-Robust algorithm operates like normal CBS, except that it also checks for collisions

which would occur if the agent is delayed by up to k steps.

Currently, the K-Robust algorithm works by creating two nodes, much like the basic CBS, each of which have time constraints which are a range (based on k).

Objective/Task

Disjoint-CBS improves basic CBS. Applying a parallel method to the K-Robust algorithm, however, proves to be trickier. This is mainly due to the positive constraint extending over a timeframe as opposed to a singular time step.

Still, we wondered if there was any way to improve upon CBS. We theorized that what causes Disjoint-CBS to improve upon basic CBS is that it creates many more constraints (a positive constraint is equivalent to a negative constraint on the rest of the agents). Based on this, we hypothesized that constraining additional agents may improve runtime. Would constraining *all* agents for each conflict improve the runtime? And if so, what would be the best method for constraining the agents?

Methodology

It has been shown [2] that choosing cardinal conflicts, conflicts which increase the solution cost, to be in child nodes reduces the overall cost/runtime of the algorithm. To show that this is true for this problem I ran some tests with the original K-Robust algorithm with the addition that for every grouping the algorithm would randomly select some number of potential groupings (5/10/25/50/100/1000) and then choose the grouping which increased the cost by the most. While the results do show that this helped slightly, it is not a practical solution as it significantly increased the runtime.

Due to this issue, I then implemented several different grouping strategies, all of which worked better than the original in select cases.

One method I tried was grouping together agents which were closest together. This entailed using an algorithm to continuously find closest points and put them in the same group until half of the agents had been placed.

I also tried the opposite - grouping together the points which were farthest away from each other.

Finally, I also tried grouping together agents whose angle direction (ie, the direction from their current location to the target) were opposite.

Results

Evaluating the modified K-Robust CBS algorithms proved challenging. Many small instances were solved too quickly by the baseline algorithm for improvements to be possible, while larger/denser scenarios often became intractable regardless of the modification. Nevertheless, several medium-complexity scenarios provided meaningful comparisons.

In these selected cases, the grouping strategies reduced the number of high-level

nodes compared to the baseline algorithm. The results are shown below.

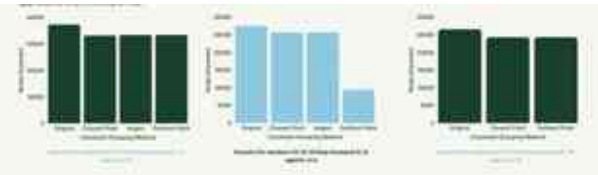


Figure 1. Results of running my algorithms on select maps/scenarios.

As the tables show, in the selected maps/scenarios all three tweaked algorithms were able to reduce the number of high-level nodes produced by ~10%.

Conclusion

These preliminary results suggest that constraining multiple agents simultaneously can improve K-Robust CBS performance, particularly when grouping is based on spatial dispersion.

The improved performance of the algorithm which grouped agents based on farthest points, especially on scenario 9, deserves a closer look – why is it that this algorithm more significantly improved node creation and runtime than the other techniques? What is it about grouping based on spatial dispersion which improves the algorithm? An answer to this question may give way to even better methods to reduce node creation and runtime of the K-Robust MAPF algorithm.

Future research could expand on these observations in two ways. First, by systematically testing across all types of maps and scenarios, the robustness of these improvements can be better established in medium-complexity scenarios. Second, more advanced approaches such as machine learning could be used to efficiently detect patterns of which groupings are preferable.

1. Li, J et al. Disjoint Splitting for Multi-Agent Path Finding with Conflict-Based Search. ICAPS (2019): 279-283.
2. Boyarski, E et al. ICBS: Improved Conflict-Based Search Algorithm for Multi-Agent Pathfinding. IJCAI (2015): 740-746.

Distributed Computing Simulator

Tani Diament
Advised under Prof. Ran Gelles

Distributed Computing is the division of a computational task across multiple nodes. In real-world scenarios, distributed algorithms are run across multiple separate computers. The challenge with this is that testing these algorithms would be very difficult and costly if it had to be done with all the physical hardware needed to test these algorithms on a large scale. This project attempts to solve that problem by letting the user simulate, visualize, and analyze the algorithm's behavior on a large set of nodes on various network topologies without the need for an actual cluster of computers. The system also lets the user test different failure scenarios such as message corruption and message loss.

The simulator is a system written in Python that simulates all the parts of a distributed system. There is a simulated computer module, a network communication module

with broadcast and directed message schemes, and an inter-computer message module. The user inputs a network topology and algorithm in a system-specific format, along with other configurations, including whether the algorithm should run synchronously or asynchronously. A synchronous algorithm runs in defined steps where everything happens in a predetermined order. Asynchronous algorithms do not have preset steps and may have unpredictable run time behavior or outcomes. There is an option to inject message corruption to model how your algorithm will behave under adverse conditions. The user can choose to run the simulator with the GUI, for up to 500 nodes, or just with an output to a text file. The simulator then runs and displays the output for the user to analyze.

The GUI allows the user to configure the system, then run the algorithm step by step, with each round being displayed in the GUI, see Figure 1. The system allows the user to skip ahead by one step, five steps, or to run the algorithm to the end at a pace chosen with a slider.

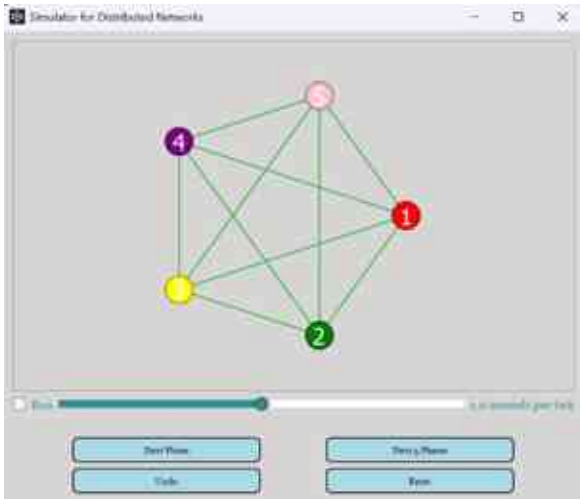


Figure 1. Simulator running leader election

The primary aspect of my work was writing distributed algorithms and setting up an automated framework to test the system. One of the distributed algorithms I wrote to test the system was a leader election algorithm based on Apache Zookeeper [1]. This algorithm helps a distributed system elect a leader on boot or after a leader. The system does so based on which node in the system has the most current information. If multiple nodes have up to date information the highest ID wins. I used a simplified form of the leader election for testing where we only look at the ID. This type of algorithm is necessary in any system where there are many computers working together to complete a task. One of the most common examples is the master-worker scheme in which the master, elected via leader election, hands out the incoming tasks to the various workers. In this type of system, the algorithm would run when the system starts to establish the master. Additionally, any time the master fails a new election would be held to find the new master which

would continue from where the old one left off. This type of scheme is used in many modern systems including Raft, Paxos, Kubernetes, MapReduce, Mesos, and many serverless platforms.

I also wrote a distributed algorithm to find the distance from any node in a network to any other node using a gossip scheme. This works in a system where each node has a vector where each index in the vector corresponds to the distance to a different node ID. The system starts with each index being distance equals infinity and the distance to yourself being zero. During each round of the algorithm every node sends their distance vector to every neighboring node. Each node then goes through the vectors and if any of its neighbors are closer to any other node you set your distance to that node as your neighbor's distance plus one. At the end each vector will have the distance to each node in the most efficient way. This type of algorithm can be used in networks to find optimal routes for message routing. This is done by sending the message to the neighbor that is closest to the node you want to send the message to. This continues until it reaches the destination node. This type of algorithm is used in real systems like RIP (Routing Information Protocol), for internet routing and BGP (Border Gateway Protocol), also used for internet routing, uses a similar algorithm.

I created GitHub Actions which automatically tests the system anytime code is pushed to the repository. This ensures that even as the project evolves, the simulator remains reliable and consistent. The automated pipeline runs the simulation on a range of test cases, covering different network topologies, algorithm types, and system configurations. By doing so, it validates that new changes do not introduce regressions, performance bottlenecks, or cause simulations to fail unexpectedly. GitHub Actions provides immediate feedback to developers by running these checks in parallel and reporting results directly on the pull request. This makes collaboration smoother, since contributors can be confident that their changes will integrate safely into the system.

We are releasing this open-source project to the public and will continue development in the GUI and feature set. The community can create pull requests to enhance the system to be reviewed by the project owners.

1. Apache Zookeeper Internals Documentation, Apache Software Foundation.

Gene Interaction Analysis Using Abstract Boolean Algebra

Andrew Haller and Moshe Wieder
Advised under Prof. Hillel Kugler and MSc student Yuval Gerber

Understanding the interactions of genes is critical to studying biological systems. A comprehensive understanding of these interactions would enable us to account for irregularities and differences in systems and thereby allow us to make further advancements in medicine and related

fields. Scientists model these interactions using Boolean Networks, where genes are either on (1) or off (0), and interactions between them are either activating (positive) or inhibiting (negative) (Figure 1). However, finding these interactions can prove difficult and require long and expensive simulations. One solution to this issue is utilizing the RE:IN (Reasoning Engine for Interaction Networks) software, which leverages Boolean algebra to produce realistic models for gene interactions [1] . First, known interactions are found from observed experiments and accepted scientific literature. To address uncertain data, Abstract Boolean Networks (ABNs) allow for optional interactions, which may or may not exist. When experimental observations are added to the model to constrain gene behavior, the result is a Constrained ABN (cABN). RE:IN can be used to model and analyze these regulatory networks and uses the Z3 SMT solver to check the satisfiability of Boolean models under experimental constraints. Each RE:IN input file defines the components (genes), regulation conditions (update functions), interactions (positive or negative; definite or optional), and the experimental conditions (specific on/off states of genes) (Figure 2).

Our modeling focuses on Boolean expressions composed only of positive literals, reflecting real biological constraints where genes are either on or off. We encode regulatory logic using formulas of the form: $(A \wedge B) \vee C$, or more generally, $(X_1 \wedge X_2 \wedge \ldots \wedge k)$ $(Y_1 \wedge Y_2 \wedge \ldots)$, which represent activation conditions over different regulatory groups [2]. Although each clause contains only positive variables, determining whether such formulas are satisfiable (consistent with biological data) is an NP-hard problem due to the combinatorial nature of the constraints across the network. These k-concatenated positive Boolean formulas allow us to model complex gene behavior while still enabling powerful logical inference using RE:IN and formal reasoning techniques. We used the RE:IN framework to assess the satisfiability of gene regulatory networks under logical constraints, allowing us to identify valid regulatory models. We saw how our gene-level constraints extended through the network by encoding regulatory interactions as structured Boolean formulas and analyzed their effect on the global behavior of the network. To represent more complex gene behavior, we built Boolean formulas with fixed-size K-groupings, making it easier to model and test different regulation patterns.

While RE:IN produced models, we extended the analysis by applying an enhanced Python implementation of RE:IN, which performs deeper verification and provides more definitive conclusions about network satisfiability and the influence of constraints on system behavior. This implementation allows for interactions to be of varying strengths and weaknesses. For example, gene X may be turned on either by gene Y being on or by genes A and B being on. A

single gene turning it on would be defined as a strong interaction, while if multiple interactions are needed, they would each be defined as weak. This real-world possibility of varying strengths can be modeled by the enhanced Python code, but not by the original RE:IN. Additionally, we used Docker to run the experiments, which made it easy to repeat or adjust the tests and get reliable results on different datasets. Docker was particularly helpful in improving the runtime of simulations and allowing for multiple simulations to run simultaneously. We made minor edits to the Python code to allow it to run in a Docker environment.

In this project, we used logic-based modeling and automation to analyze complex gene regulatory networks. By combining the RE:IN tool with the Python code, we were able to test and verify which regulatory models are consistent with given constraints. While RE:IN helped generate possible models, the Python implementation provided stronger verification and clearer results. Running everything in a Docker environment ensured that our experiments were consistent and easy to reproduce. This approach helps pave the way for future studies on larger scales in computational biology and gene network modeling.

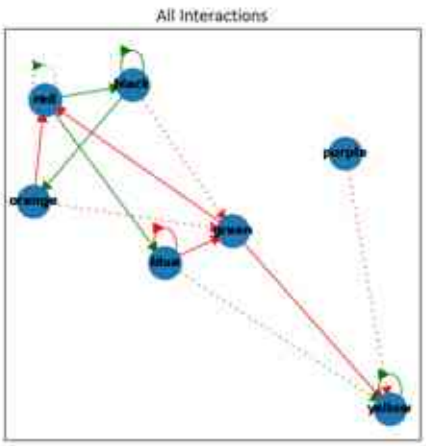


Figure 1. A model of interactions represented by the RE:IN software. Green lines represent positive interactions, while red lines indicate negative interactions. Solid lines represent definitive interactions, while dotted lines are optional interactions.



Figure 2. A sample RE:IN file containing components A-Y with all regulation conditions, component interactions consisting of optional and verified interactions along with interaction of varying strengths, conditions, and experimental conditions.

1. Yordanov, B., Dunn, S.-J., Kugler, H., Smith, A., Martello, G., & Emmott, S. (2016). *A method to identify and analyze biological programs through*

automated reasoning. npj Systems Biology and Applications, **2**, 16010.

2. Y. Gerber. Reasoning about Monotonic Regulation Conditions for Boolean Gene Regulatory Networks: An Algorithmic Approach, 1-14.

Optically Identifying Bacteria

Aviva Klahr
Advised under Prof. Dror Fixler, Dr. Hamootal Duadi, MSc student Bar Atuar

Detecting and characterizing bacterial contamination in water is essential for maintaining public health and environmental safety. Conventional microbiological and biochemical techniques, such as membrane filtration, enzyme–substrate assays, and molecular diagnostics, offer high accuracy and specificity but are often slow, costly, and equipment-intensive. These methods typically require culturing, reagents, and specialized handling, limiting their usefulness for rapid or large-scale monitoring. Our project explores an alternative, optical approach: identifying bacterial contamination based solely on how light interacts with water samples, specifically through light scattering and absorption properties. This optical method has several advantages: it allows for rapid, real-time measurements without the need for culturing or chemical processing, it can be implemented using low-cost and reusable components such as LEDs, lasers, and photodetectors, and it enables repeatable, non-destructive testing suitable for continuous monitoring.

Light interacts with matter primarily through two mechanisms: absorption and scattering. Absorption occurs when photons are absorbed by molecules in the medium, transferring their energy to those particles and reducing the transmitted light intensity. Each substance absorbs light at specific wavelengths, producing a characteristic absorbance spectrum that can be used for identification. Scattering, by contrast, occurs when particles in the medium deflect photons from their straight trajectory, redirecting them in multiple directions. In turbid materials or suspensions such as bacterial cultures, both effects occur simultaneously, reducing total transmitted light intensity. Traditional spectrophotometers, which rely on the Beer–Lambert law, measure total light attenuation and interpret all lost intensity as absorption, effectively conflating the effects of absorption and scattering. This makes them unsuitable for analyzing turbid samples where both phenomena contribute significantly to light loss.

To address this limitation, our lab employs a novel optical technique based on the discovery of an *iso-pathlength point* (IPL). The IPL is defined as the specific detection angle at which the measured light intensity becomes independent of the medium's scattering properties, depending solely on absorption. This unique property allows the IPL to serve as a natural calibration point and a tool for differentiating scattering and absorption effects in complex media.

The experimental setup (Figure 1) consists of a circular glass dish containing the liquid sample. A fixed light source, composed of both white light and a blue LED, illuminates the sample, providing a broad spectrum of wavelengths. On the opposite side, an optical fiber connected to a spectrometer measures the transmitted and scattered light intensity. This detector is mounted on a step motor that rotates around the dish's circumference, allowing precise measurements at varying angles relative to the light source. The system was carefully designed and refined for repeatability, including the use of 3D-printed mounts to ensure consistent sample positioning.

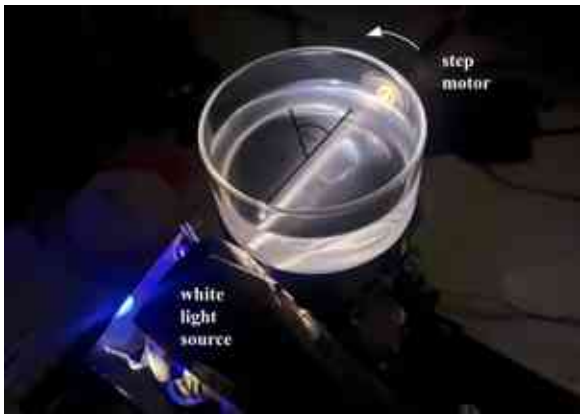


Figure 1. Experimental setup

To characterize the IPL, we first tested the system using titanium dioxide (TiO_2) suspensions of varying concentrations as controlled scattering samples. Starting with double-distilled water (DDW), which exhibits negligible scattering and absorption, we incrementally increased TiO_2 concentration, observing how intensity varied with angle. As expected, at small angles, higher scattering concentrations resulted in lower intensities because photons were deviated from the direct path. However, as the detection angle increased, this relationship reversed: more scattering led to higher measured intensity because scattered photons were redirected toward the detector. Remarkably, the intensity–angle curves for all concentrations intersected at a single point, approximately 2.8°, indicating the existence of an angle where intensity is independent of scattering (Figure 2). This intersection defines the IPL. At this angle, any reduction in measured intensity arises solely from absorption, not from scattering. Thus, the IPL allows direct measurement of absorption even in highly scattering media.

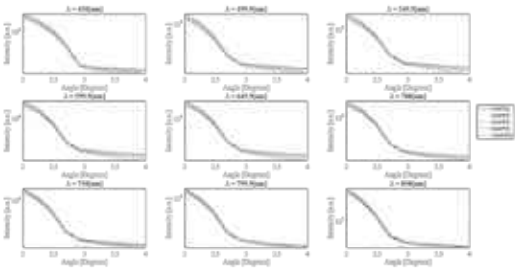


Figure 2. IPL seen with TiO_2 samples: Intensity vs. angle at specified wavelengths

Having established the IPL with scattering materials, we extended the technique to biological samples, specifically three bacterial species commonly found in water: *Pseudomonas putida*, *Escherichia coli*, and *Listeria monocytogenes.* The goal was to determine whether differences in each bacterium's optical scattering and absorption spectra could serve as

identifying signatures. For each bacterial suspension, we measured the intensity spectra I/I_{IPL} versus wavelength at multiple angles. Individually, the bacterial curves appeared similar, but when plotted together (Figure 3), distinct spectral differences emerged. *P. putida* exhibited stronger scattering in the green region (~500 nm), while *Listeria* showed a peak in the orange–red range (~580–630 nm). *E. coli* displayed an intermediate pattern. These observations suggest that different bacterial species scatter light preferentially at distinct wavelengths, providing a potential basis for spectral identification.

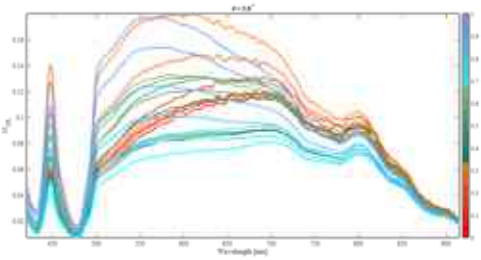


Figure 3. EC, PP, Lis: I/I_{IPL} vs. wavelength at specified angle

To enhance these distinctions, we further normalized each bacterial spectrum by dividing I/I_{IPL} by the DDW reference spectrum (Figure 4). This produced clearer, species-specific curve slopes, indicating that each bacterium possesses a unique scattering–absorption "fingerprint." Although more quantitative analysis is ongoing, these patterns already demonstrate the feasibility of distinguishing bacterial species in water samples using purely optical data.

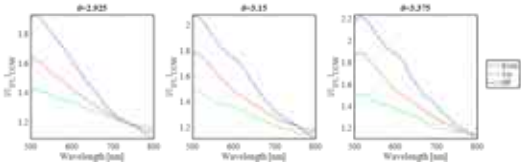


Figure 4. EC, PP, Lis at certain concentration: $I/I_{IPL}/I_{DDW}$ vs. wavelength at specified angles

Our results lead to several important conclusions. First, the IPL is a genuine physical phenomenon that can be experimentally identified: an angle at which light intensity depends only on sample geometry and absorption, and not on scattering. Second, the IPL provides a practical means to decouple absorption from scattering, something traditional spectrophotometric techniques cannot achieve. This capability offers a new, non-invasive approach to characterizing turbid or particulate samples. Finally, preliminary bacterial measurements suggest that this technique could be extended to real-world applications in water quality monitoring. By leveraging the unique optical signatures of bacterial species, it may become possible to rapidly identify and quantify bacterial contamination without culturing, reagents, or complex instrumentation. While further refinement and statistical analysis are needed, the results demonstrate the potential for a fast, cost-effective, and non-destructive optical biosensing technique grounded in fundamental light–matter interactions. This work lays the foundation for developing automated, real-time water quality sensors that could have significant implications for

environmental monitoring, public health, and industrial applications.

In conclusion, the iso-pathlength point offers a transformative framework for optical analysis of complex media. By isolating absorption effects from scattering, it enables new methods for distinguishing bacterial species and monitoring contamination in water. While further refinement and statistical analysis are needed, the results demonstrate the potential for a fast, cost-effective, and non-destructive optical biosensing technique grounded in fundamental light–matter interactions. This work lays the foundation for developing automated, real-time water quality sensors that could have significant implications for environmental monitoring, public health, and industrial applications.

Detection of Atto 532 Fluorescent Dye Using a High Throughput Optical Modulation Biosensing System

Rachel Sharon
Advised under Prof. Amos Danielli and PhD student Shmuel Burg

Early and accurate detection of low-abundance biomarkers is fundamental for improving patient outcomes, particularly in the context of diagnosing infectious diseases, cancer, and other conditions. Conventional immunoassays such as enzyme-linked immunosorbent assays (ELISA) remain widely used but often lack the capacity to detect very small quantities of biomarkers. While magnetic bead-based assays have improved sensitivity by providing a platform for capturing biomolecules, they are still limited by their slowness, being labor-intensive, and having high costs. As a result, there is a growing need for diagnostic technologies that combine high sensitivity with rapid turnaround. The high-throughput optical modulation biosensing (ht-OMBi) system [1] integrates the sensitivity of magnetic bead-based assays with an optical detection strategy that minimizes background noise and accelerates data acquisition. ELISA, the gold standard assay, obtains a fluorescent signal from the entire well, whereas in the ht-OMBi system, the fluorescent signal is concentrated into the center of the well by the magnetic beads, which allows for the increased sensitivity of this assay.

In the ht-OMBi system, a sandwich assay (Figure 1) is used, in which magnetic beads capture fluorescently tagged target molecules in the well and are aggregated using a conical magnet tip. A laser beam is then directed at the well and moves between regions with and without the aggregated beads. By comparing the fluorescence signal during these modulated movements, the system effectively reduces noise from background fluorescence. The system can read a full 96-well plate in under 10 minutes, offering faster throughput than ELISA or qPCR [2]. Previous studies have validated its applicability to diverse biomarker targets, including proteins (IL-8, ricin), viral antigens (SARS-CoV-2, Zika), antibodies (anti-dengue NS1 IgG), and bacterial components (Streptococcus A).

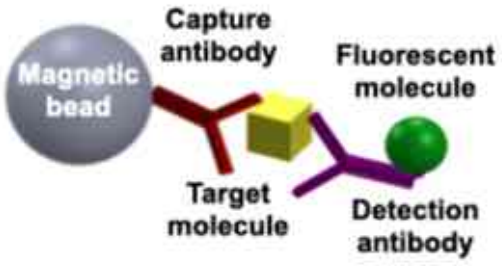


Figure 1. Biological sandwich assay

In these experiments, we sought to calibrate the ht-OMBi system using Atto 532, a fluorescent dye. Our objectives were threefold: (1) establish the relationship between Atto dye concentration and fluorescence signal across different OMBi systems, (2) evaluate the effect of photobleaching [3] and treatment on bead autofluorescence by comparing treated and untreated magnetic beads, (3) assess well-to-well variability to determine reproducibility and accuracy of the platform.

The experimental procedure began with the preparation of assay and reading buffers, followed by a ten-fold dilution series of Atto 532 solutions. Beads were photobleached for 20 hours to minimize autofluorescence and subsequently divided into ten tubes (100,000 beads per tube). Each tube received a distinct Atto concentration, spanning from 0 molecules per bead (negative control) up to 10^7 molecules per bead. If the assay required naked beads, no Atto dye solution was added. Following incubation with the dye, beads were magnetically separated, washed twice, and resuspended in reading buffer. For assays involving untreated naked beads, the washing steps were skipped. For plate-based analysis, beads were transferred to black 96-well plates at 25,000 beads per well, with each concentration distributed across four replicate wells. The fluorescence signals were then recorded three times per plate using three separate OMBi systems. The data was then analyzed, entered into graphs or charts, and the Limit of Detection (LoD) and coefficient of variation (CV) were calculated.

The results in Figure 2 demonstrate a dose-response relationship between Atto 532 concentration and fluorescence intensity in all three OMBi systems. Systems 1 and 3 displayed highly consistent signal profiles, characterized by similar LoDs and CVs. In contrast, System 2 exhibited a slightly elevated limit of detection and higher CV. These findings confirm the reproducibility of OMBi calibration across multiple setups. Figure 3 compares naked beads that were treated with the washing assay versus untreated naked beads. We found that washed beads exhibited approximately double the average fluorescence intensity and half the CV relative to unwashed beads. Importantly, the standard deviations of treated and untreated groups were nearly identical. Further analysis of plate-level reproducibility was conducted by calculating average fluorescence, maximum fluorescence, standard deviation, and CV across all wells. As seen in Figure 4, heatmaps of well-to-well variability demonstrated that average fluorescence

intensities were generally uniform across the 96-well plate, typically ranging between 800 and 1200 units. Moreover, CV values were consistently low, with most wells falling below 20%. These findings confirm that the ht-OMBi platform provides reliable and reproducible measurements across large sample sets.

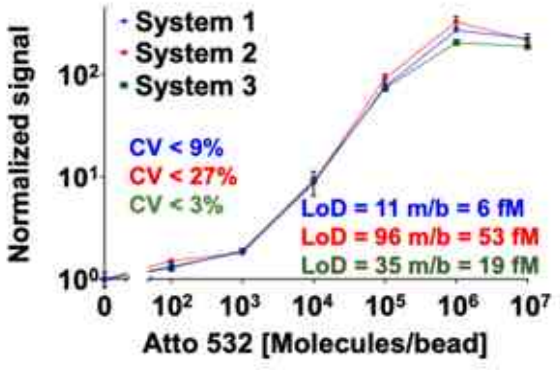


Figure 2. Atto 532 dose response in 3 OMBi systems

Stat	Untreated (no assay)	Treated (assay)
AVG	1078	1960
SD	209	201
CV	19	10

Figure 3. AVG, SD, CV of untreated vs treated naked beads

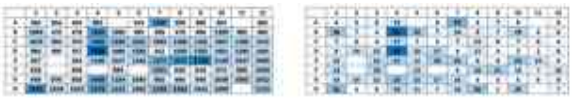


Figure 4: 96-well plate heatmap, average value (left) & CV (right)

In conclusion, these results validate the utility of the ht-OMBi system for sensitive, rapid, and reproducible fluorescence-based biosensing. Calibration with Atto 532 dye not only establishes a benchmark for system performance but also highlights key parameters that influence assay optimization. The findings suggest that ht-OMBi could significantly advance diagnostic testing by combining the sensitivity of magnetic bead-based assays with the speed and scalability of optical modulation. Ultimately, the ht-OMBi system holds promise for transforming biomarker detection in both research and clinical diagnostics by enabling early, accurate, and high-throughput analysis of low-abundance targets.

[1] S. Burg et al., High throughput optical modulation biosensing for highly sensitive and rapid detection of biomarkers, Talanta **248**, 123624 (2022).
[2] S. Burg et al., From Concept to Commercialization: High-Throughput Optical Modulation Biosensing for Detecting Low Concentrations of Biomarkers, in S. Karakus, Current Developments in Biosensor Applications and Smart Strategies, IntechOpen, (2025).
[3] S. Roth et al., Improving the sensitivity of fluorescence-based immunoassays by photobleaching the autofluorescence of magnetic Beads," Small **15**, 1803751 (2019).

Testing Natural Oscillator Synchronization on an FPGA

Anna Weisman
Advised under Dr. Itamar Levi

When two or more oscillators are placed close together in a circuit, they can naturally synchronize due to coupling effects, even if no explicit wired connection exists between them. This phenomenon occurs because oscillators generate time-varying electrical and magnetic fields that can influence nearby oscillators through parasitic

capacitance, mutual inductance, or electromagnetic radiation. In integrated circuits, additional coupling can occur through the shared power supply or the silicon substrate. If the oscillators start at frequencies that are close to one another, these weak interactions can gradually pull them into frequency and sometimes phase alignment. The strength of the coupling determines whether the oscillators only shift slightly toward each other, a behavior known as frequency pulling, or whether they fully lock together. Natural synchronization can have both positive and negative consequences: in some RF systems, it can reduce phase noise, but in other applications such as random number generators that rely on independent oscillators, it can cause serious performance issues. Because of this, circuit designers often need to carefully consider layout, spacing, and isolation techniques to either promote or prevent synchronization depending on the application.

In our case, we want to take advantage of the coupling and transport phase information across a chip. If there is a chain of oscillators with known synchronization parameters, we believe that it is possible to transport information along this path. The benefits of this method may include low power clock synchronization, or an added layer of encryption. This project is progress towards setting up an environment to test the parameters of the synchronization of an oscillator chain on a field programmable gate array (FPGA). The idea is to set up ring oscillators physically close together on the FPGA. We want to control the starting phases of the oscillators and measure how the oscillators react to their neighbors. Since we want to set up many tests, our design must be parametrized. The details of each test should be communicated to the FPGA from a computer without having to re-synthesize and re-program the FPGA between trials. To do this we will communicate with the FPGA via Universal Asynchronous Receiver-Transmitter (UART) using pySerial. We want the options to change the number of stages in each oscillator, the number of oscillators, the distance between them, the phase shifts at the beginning of the testing, and the length of time we want to run the experiment for.

The vision is to employ a top level system (written in Verilog) which will control the testing environment. In the Idle state, the FPGA will produce a signal saying that it is ready to receive commands. Once the computer sends a signal that it is indeed sending information, the FPGA will listen and enter a Setup mode. In this mode, the FPGA will configure the correct testing environment and then move into the Run state. In the Run state, the FPGA will produce a trigger signal for the external oscilloscope. This is to listen to the FPGA from the beginning of its experiment. This special attention to the beginning of the experiment is necessary since we are focusing on the behavior of the phases of the oscillators before they fully synchronize. After running until the specified number of clock cycles, the FPGA will move to the Send state. In the send state, the FPGA will send a

message to the computer. Figure 1 shows a diagram of this finite state machine.

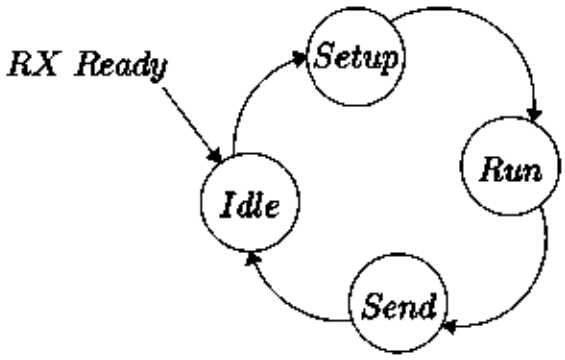


Figure 1. The finite state machine depiction of the system

The first step to implementing the testing environment is to implement ring oscillators on the FPGA. A ring oscillator is a circuit consisting of a chain of an odd number of not gates. Our goal is to measure the rate of synchronization of the oscillators. In order to do this, we need the oscillators to start off with a known phase difference that we can study. To arrange for this, we want to control the enable signals to produce time delays in the starts of each oscillator. The value of the phase delay should be instructed by the user at the computer. Meaning, this parameter should be sent via UART to the FPGA. It will be useful in the future to be able to specify the phases of each oscillator, however, as a starting point, we will use a single delay value to offset all of the oscillators. The delay value will be the amount of delay between the start of one oscillator to the next. Figure 2 shows the GTKWave (an application to view waveforms) results of the controlled oscillations of the parametrized Verilog module.

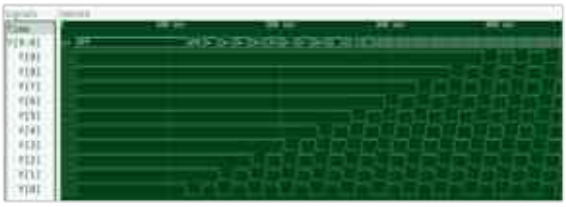


Figure 2. Simulation results of the parametrized Verilog Module

Since we want to test this system at a high frequency of 1 GHz, we need to find a viable way of measuring this signal. Measuring a low-voltage, high-frequency signal such as a 1 GHz waveform with an oscilloscope is challenging because the act of probing disturbs the signal itself. At such high frequencies, the probe and oscilloscope are no longer just “observers,” but part of the circuit. The main issue comes from the probe’s finite input capacitance and resistance in its leads. This extra load can significantly attenuate the signal, distort its shape, or even slow down the edge transitions. Since the signal is already low voltage, any attenuation or distortion introduced by the probe becomes more severe.

One method we may employ to bypass the limitations of our equipment is to slow down our signals by a known amount. Since we are only interested in measuring the phases of the waveforms, it does not bother us if all the waveforms were slowed down by the same multiple. This is because phase is relative to the original waveform. If both signals are slowed down by the same factor,

we can still read an accurate value for our phase measurement. To do this we can pass our signals through a frequency splitter shown in Figure 3.

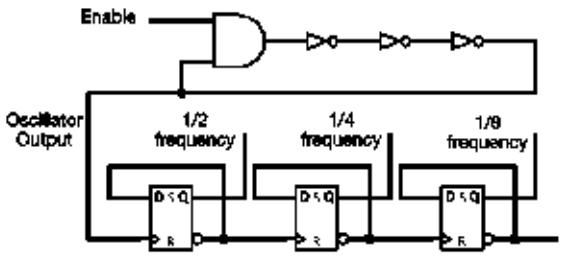


Figure 3. The structure of the oscillator and frequency splitter circuit.

The Spartan6 FPGA device Sakura-G was used to set up a testing environment to verify and measure the effects of natural oscillator synchronization due to proximity. The environment takes commands from a computer and runs the desired experiment. The results are monitored on an oscilloscope. Future work includes testing the UART system, setting up the experiment, and measuring the results. The hope is that this system could be used for further research in the field of natural oscillator synchronization, and after measuring the parameters and creating a model to describe the phenomena, we could design a new method of signal transportation. This new method may offer benefits in the areas of power consumption, encryption, or conserving space.

LIFE SCIENCES



Rachel Schwartz, Ariel Melnitsky, Anna Laufer, Talia Hazan, Ahuva Halpert, Leah Sher, Ben Epstein

A Study of SIRT6 and the Molecular Biology of Aging

Ben Epstein and Leah Weiss
Advised under Prof. Haim Cohen and PhD students Ron Nagar and Zacharia Schwartz

Biomedical research frequently aims at one goal: to uncover groundbreaking mechanisms and therapies for various diseases. The focus of this project, however, is to study a universal phenomenon – aging. This biological process is perhaps the underlying root of many illnesses later in life, thus rendering it a worthwhile process to study. Scientists have long strived to increase healthspan and lifespan to ensure a greater length of life in good health.

Some of the most significant hallmarks of aging include changes in chromatin structure and gene expression. These changes lead to various adverse effects such as increased inflammation, metabolic decline, and tissue dysfunction. Among the molecular regulators involved in these processes is Sirtuin6 (SIRT6), one of the seven mammalian sirtuins. SIRT6 has several functions such as regulating metabolism and modulating inflammation, as well as helping to maintain genomic

integrity. Additionally, as an NAD+-dependent enzyme, SIRT6 functions as a histone deacetylase; SIRT6 removes acetyl groups from histone H3 lysine residues, which keeps chromatin in its condensed state [1]. However, during aging, SIRT6 levels decrease. The histones then remain acetylated, prompting the loosening and increased accessibility of chromatin. Chromatin alterations accompany transcriptional changes, contributing to the loss of each cell's specialized identity.

Previous experimentation has shown that although a small subset of chromatin regions becomes more closed with age, the overall trend is increased chromatin accessibility. Regions that close are related to β-oxidation and liver function, whereas the regions that open are associated with immune system activation. This aligns with the physiological effects of aging seen in previous studies, including a decline in liver function and increase in inflammation.

This project focuses on identifying transcription factors that facilitate the expression of genes in age-associated open chromatin regions, which are primarily linked to inflammation. In light of their strong association with inflammation, the ETS family of transcription factors were selected for study to pinpoint specific members that drive age-associated inflammation, in order to better understand the precise mechanisms by which SIRT6 maintains healthspan and promotes longevity. Supporting analyses were performed to assess whether SIRT6 functions independently of other histone deacetylases, and to determine its role in repressing retrotransposable element LINE1 expression.

To identify members of the ETS family that are upregulated with age, and determine whether SIRT6 overexpression can mitigate these age-associated changes, quantitative PCR (qPCR) for several ETS genes was performed on cDNA prepared from liver tissue of wild-type and transgenic mice overexpressing SIRT-6, raised to "young" and "old" ages (5-7 months and 18-21 months, respectively).

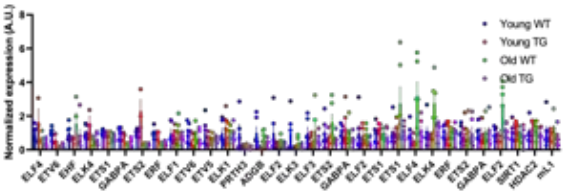


Figure 1. qPCR analysis of gene expression across all transcription factors and targets studied

Of the 15 ETS genes tested in whole liver tissue (Figure 1), ERF expression was shown to be significantly upregulated with age in wild-type mice, but reversed by SIRT6 in transgenic mice (Figure 2). This suggests its role in age-related inflammation and that ERF may play a role in causing transcriptional changes in aging.

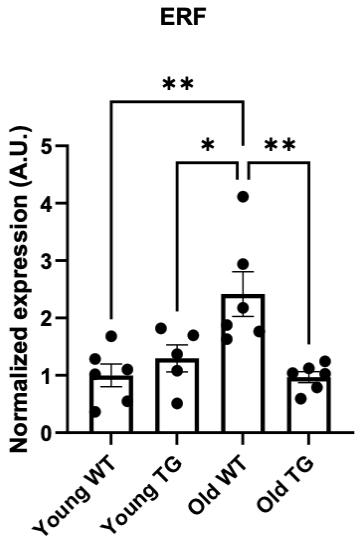


Figure 2. qPCR analysis of ERF expression in whole liver tissue. Statistical significance: *, $p < 0.05$, **, $p < 0.01$, ***, $p < 0.001$. Outliers were excluded.

However, when qPCR was performed on isolated hepatocytes, ERF expression levels were higher in young mice than old mice, with minimal differences between the wild-type and transgenic groups (Figure 3). This suggests that the trend seen in whole liver tissue does not hold in isolated hepatocytes, indicating that the effects of SIRT6 are seen in non-hepatocyte cells in the liver.

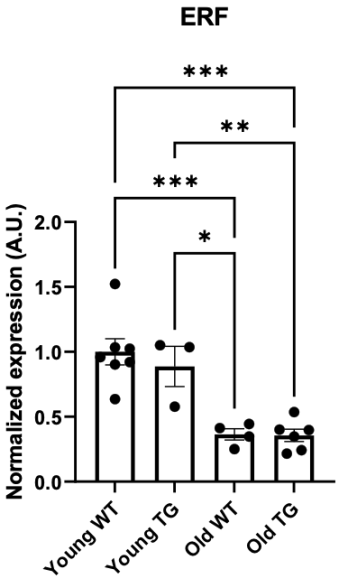


Figure 3. qPCR analysis of ERF expression in hepatocytes. Statistical significance: *, $p < 0.05$, **, $p < 0.01$, ***, $p < 0.001$. Outliers were excluded.

To confirm that changes in chromatin and gene expression were due to the activity of SIRT6 and not other histone deacetylases, expression of HDAC2 was tested across all groups. HDAC2 levels remained essentially unchanged across all groups, supporting that SIRT6 acts directly and not through activation of other deacetylases (Figure 4).

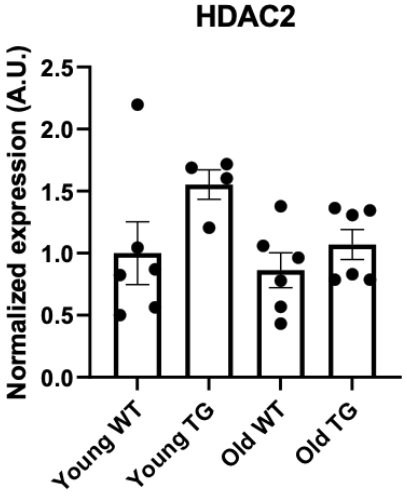


Figure 4. qPCR analysis of HDAC2 expression in whole liver tissue. Statistical significance: *, $p < 0.05$, **, $p < 0.01$, ***, $p < 0.001$. Outliers were excluded from analysis.

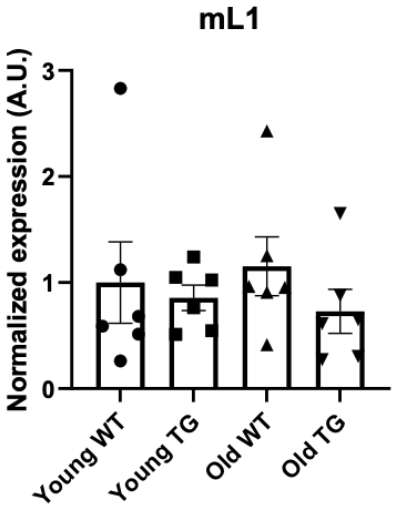


Figure 5. qPCR analysis of LINE1 expression in whole liver tissue. Statistical significance: *, $p < 0.05$, **, $p < 0.01$, ***, $p < 0.001$. Outliers were excluded from analysis.

The expression levels of LINE1, a retrotransposable element generally located in closed chromatin, were tested in liver tissue. During aging, these chromatin regions tend to become more open, allowing for LINE1 expression and, therefore, inflammatory response. qPCR results showed lower LINE1 expression in old TG mice compared to old WT mice, though an insignificant trend (Figure 5). The results reveal that SIRT6 may help suppress LINE1 expression and activity, helping to maintain chromatin integrity.

With 30 genes studied as well as promising results found, this project contains important information to understanding the molecular biology of longevity and healthspan. Further experimentation and qPCR analysis is essential in deepening our understanding of SIRT6 as an anti-aging protein.

1. M. Baran, P. Miziak, A. Stepulak, M. Cybulski, The role of Sirtuin 6 in the deacetylation of histone proteins as a factor in the progression of neoplastic disease. Int. J. Mol. Sci. **25**(1), 497 (2023).

Tracking the Stress Granule Dynamics of the RNA-Associated Protein MOV10

Ahuva Halpert
Advised under Prof. Yaron Shav-Tal and MSc student Koren Walk

Stress granules and P-bodies are cytoplasmic RNA-protein granules that play

crucial roles in post-transcriptional gene regulation. P-bodies are present under normal cellular conditions and function in mRNA degradation, storage, and silencing. In contrast, stress granules are not always present but form rapidly in response to environmental stressors like oxidative damage or nutrient deprivation. Stress granules temporarily store untranslated mRNAs, pausing translation to allow the cell to focus resources on recovery and survival. Among well-characterized markers of these granules, G3BP1 is a core stress granule nucleator, essential for stress granule assembly under stress conditions. DDX6 serves as a central component of P-bodies, involved in mRNA decapping and decay and Hedls plays a role in P-bodies assembly too. Despite the known background of these proteins, the RNA helicase MOV10 remains poorly characterized in the context of RNA granule biology. MOV10 is known for its role in RNA silencing and antiviral defense. However, its potential association with stress granules or P-bodies has remained unclear. Understanding if and how MOV10 participates in granule dynamics could give new insights into post-transcriptional gene regulation during cellular stress.

The goal of this study was to track the localization and behavior of MOV10 in live cells under two distinct stress conditions: oxidative stress induced by sodium arsenite, which promotes stress granule assembly, and amino acid deprivation induced by EBSS which enhances P-bodies. By using established markers for stress granules (G3BP1) and P-bodies (DDX6), we tried to determine whether MOV10 associates with one granule type, dynamically transitions between them, or remains none under stress.

Methodology

This study aimed to examine how the RNA-associated protein MOV10 and the stress granule marker DDX6 respond to different types of cellular stress. We started by looking at untreated cells to establish a baseline for protein localization. Cells were then exposed to nutrient stress using EBSS and oxidative stress using arsenite to induce stress granule formation. Immunofluorescence (IF) was used to visualize protein localization, with fluorophore-conjugated secondary antibodies binding to primary antibodies that recognize the proteins of interest. Western blotting was performed for DDX6 to measure protein levels, separating proteins by size on a gel, transferring them to a membrane, and probing with specific antibodies. RNA FISH was also used to see where DDX6 RNA was in the cells, using fluorescently labeled probes that bind to target RNA. IF was performed under both stress conditions to observe the localization of MOV10.

Results

The distribution of P-body and stress granule proteins was studied in U2OS cells under normal and stress conditions using immunofluorescence and western blotting. In Figure 1, untreated cells are shown. DDX6 staining marks the presence of P-bodies, while G3BP1 staining demonstrates that no

stress granules are present under baseline conditions. This confirms that in the absence of stress, P-bodies are visible, but stress granules remain dispersed in the cytoplasm.

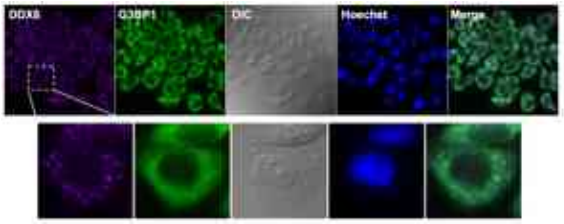


Figure 1. U2OS cells before induced stress; P-bodies are seen with the DDX6 protein. G3BP1 protein show SG remains dispersed in cytoplasm.

Next, arsenite stress was applied. In Figure 2, P-bodies are again detected with DDX6 staining, and in contrast to baseline, stress granules are now visible with G3BP1. This demonstrates that arsenite induces stress granule formation alongside P-bodies. To validate that DDX6 protein levels were not altered by the treatment, a western blot was performed (Figure 3). Strong DDX6 bands were observed, confirming that stress does not reduce overall protein abundance, but instead influences protein localization.



Figure 2. U2OS cells after arsenite treatment; P-bodies are seen with the DDX6 protein and stress granules are formed.



Figure 3. Western blot of DDX6 protein

After establishing this system, we next focused on the distribution of MOV10. In Figure 4, untreated cells show MOV10 dispersed throughout the cytoplasm with no association to RNA granules. Under EBSS-induced stress, MOV10 relocalized to P-bodies (Figure 5), while under arsenite-induced stress, MOV10 redistributed to stress granules that were closely associated with P-bodies (Figure 6).

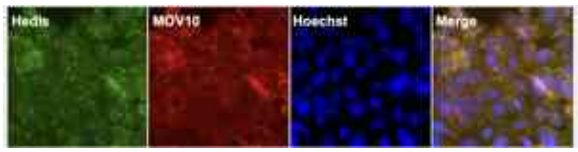


Figure 4. U2OS cells with no treatment; MOV10 dispersed in cytoplasm

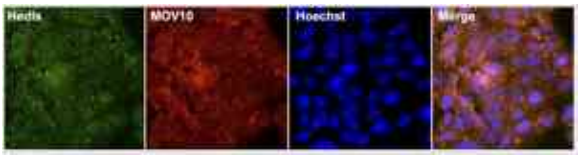


Figure 5. EBSS induced stress; MOV10 localizes to P-bodies

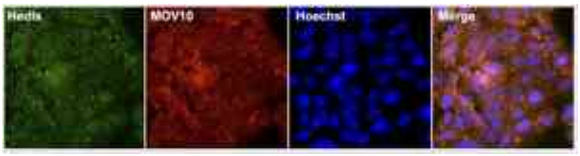


Figure 6. Arsenite induced stress; MOV10 localizes to stress granules that associate with P-bodies

Conclusion

Our results show that MOV10's localization changes depending on the type of stress the cell experiences. In untreated cells, MOV10 was spread throughout the cytoplasm and not associated with RNA granules. When cells were stressed with EBSS, MOV10 moved into P-bodies, which may help the cell store or break down mRNAs when resources are limited. Under arsenite stress, MOV10 instead moved into stress granules, which protect mRNAs from being degraded. Interestingly, under arsenite stress, MOV10 was found at stress granules that were closely associated with P-bodies. This suggests that P-bodies and stress granules may share some functions or interact with each other, a connection that could be further studied in the future. Western blot results showed that DDX6 levels stayed strong under all conditions, meaning that stress does not reduce the amount of protein but instead changes its location in the cell. This supports the idea that P-body and stress granule dynamics are controlled by relocalization rather than expression. Taken together, our findings show that MOV10 is a dynamic protein that responds to different stress signals by moving into specific RNA granules. Its presence in both P-bodies and stress granules under stress highlights a potential link between these compartments in regulating mRNA fate. Future experiments could include live-cell imaging to track MOV10 movement, RNA immunoprecipitation to identify transcripts associated with each granule type, and knockdown or overexpression studies to see how MOV10 affects granule assembly. Overall, this work suggests that stress granules and P-bodies are connected components of a network that helps cells respond to stress, and MOV10 may play an important role in coordinating their activities.

Applying Metaproteomics to Investigate the Interaction Between Cancer Tumors and the Gut Microbiome

Talia Hazan
Advised under Dr. Lior Lobel and PhD student Tal Nissenbaum

The human gut microbiome has emerged as a critical regulator of host physiology, influencing metabolism, immune responses, and susceptibility to disease. Increasing evidence has linked specific microbial communities to both cancer initiation and progression. In melanoma, a form of skin cancer, immune checkpoint blockade (ICB) therapy has transformed treatment outcomes by restoring anti-tumor immunity. Despite its success, only a subset of patients respond to ICB therapy and the biological mechanisms underlying these differences remain uncertain. Several studies have associated gut microbial composition with response outcome, yet

most relied exclusively on metagenomic sequencing, which identifies the genetic blueprint but does not reveal whether microbial genes are actively expressed. Thus, there remains a need to assess functional activity at the protein level in order to understand microbial contributions that directly influence a patient's response to immunotherapy.

To address this uncertainty, we applied a metaproteomic approach that uses mass spectrometry and a cohort-specific metagenomic database, enabling identification and quantification of microbial proteins in patient's stool samples. Stool specimens were obtained from melanoma patients who have undergone ICB therapy in Pittsburgh, Pennsylvania. Patients were separated into two groups based on their clinical response: responders, defined by partial or complete tumor regression, and non-responders, defined by disease progression despite therapy. Both proteins and DNA were extracted from all stool samples. Protein concentrations were quantified using the Qubit protein assay, standardizing each to 100 µg for downstream analysis. The samples were then prepared using the Enhanced Filter-Aided Sample Preparation (eFASP) method, which lyses, solubilizes, and digests proteins using trypsin, while removing contaminants to generate clean peptide samples for mass spectrometry. To ensure consistency across runs, pooled samples were added to each batch for statistical normalization. Peptide concentrations were quantified using a Quantitative Fluorometric Peptide Assay to verify accurate input amounts, and finally, the samples were labeled with Tandem Mass Tags (TMT) to enable multiplexed proteomic analysis and identification of peptides from each sample. The samples were subjected to multiplexed LC-MS analysis using a mass spectrometer. (Figure 1) To improve peptide identification beyond public reference databases, we constructed a customized, cohort-specific protein database from shotgun metagenomic sequencing of the same samples, capturing the unique microbial landscape of each patient cohort.

This multi-omics approach generates a dataset that significantly exceeds conventional pipelines.

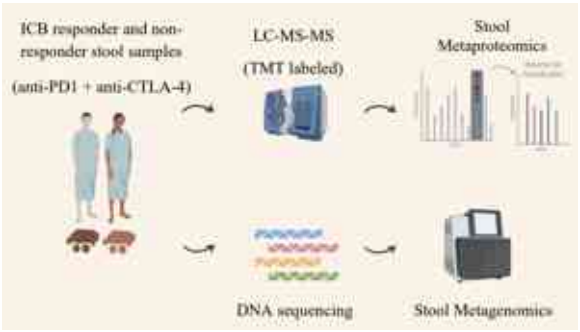


Figure 1. Workflow for metagenomic and metaproteomic analysis of stool samples from melanoma patients treated with ICB therapy.

In the previous project, which applied a parallel design using an Israeli cohort from Sheba Medical Center, the comparative analysis of responders versus non-responders revealed notable functional

differences. Within this cohort, taxonomic comparisons between metagenomic and metaproteomic data showed similar bacterial distributions, but the proteomic data revealed a greater representation of Bacteroidetes proteins and a lower relative abundance of Actinobacteria proteins, suggesting differences in microbial activity that are hidden at the genomic level (Figure 2). Over 200 bacterial proteins exhibited differential abundances. Pathway enrichment analysis and gene set enrichment analysis identified metabolic signatures that distinguished the two groups. Most notably, glycolysis-related proteins were enriched in non-responders, suggesting that bacterial utilization of host-derived glucose may contribute to local nutrient depletion, impairing T-cell function and undermining the effectiveness of immunotherapy. In addition, methanogenesis pathways, particularly those originating from CO_2 reduction, were significantly elevated in non-responders. This enrichment implies that archaeal activity may play an immunomodulatory role, potentially by producing methane with anti-inflammatory effects or by reshaping community structure to favor immune evasion. Importantly, these differences were not evident from metagenomic analysis alone, as glycolysis and methanogenesis genes are widespread across microbial taxa. Only by examining protein-level expression were these functionally relevant distinctions revealed.

These findings highlight the power of metaproteomics to uncover mechanistic insights that cannot be obtained through genomic sequencing alone. Whereas metagenomics highlights which species are present, metaproteomics reveals which functions they are carrying out in the tumor-bearing host. By capturing dynamic microbial activity, this study deepens our understanding of how the gut microbiome influences immunotherapy response at the functional level.

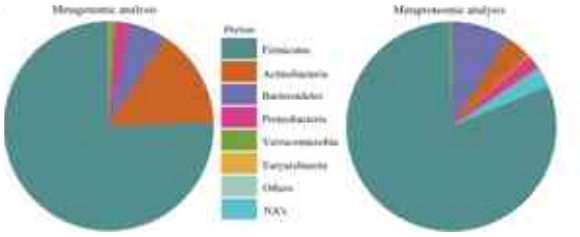


Figure 2. Taxonomic distribution of gut bacteria by phylum from metagenomic (left) and metaproteomic (right) analyses

The previous study demonstrated that metaproteomic profiling of stool samples from melanoma patients undergoing ICB therapy identifies protein-level differences between responders and non-responders. The enrichment of glycolytic enzymes and methanogenesis proteins in non-responders points to potential microbial mechanisms of resistance, including metabolic competition with host immune cells and the generation of immunosuppressive metabolites. These results not only validate the utility of metaproteomics in cancer microbiome research but also pave the way toward developing microbiome-derived biomarkers for patient stratification. We aim to

replicate these findings in the Pittsburgh melanoma cohort to validate and strengthen the observed associations between microbial protein activity and immunotherapy response. Ultimately, the goal is to apply this approach *in vivo* using mice models to investigate causal relationships between specific microbial functions and tumor progression. Long-term, this research seeks to identify microbiome-derived biomarkers that could facilitate early cancer detection and advance therapeutic strategies to modulate the tumor microenvironment. By integrating functional microbiome data into clinical oncology we move closer to precision strategies that harness the microbiota as both a predictive marker and a therapeutic partner in the fight against cancer.

1. Gopalakrishnan V., Spencer C.N., Nezi L., Reuben A., Andrews M.C., *et al. Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients.* Science; **359**(6371):97-103 [2018]

2. Gopalakrishnan V, Helmink BA, Spencer CN, Reuben A, Wargo JA. The Influence of the Gut Microbiome on Cancer, Immunity, and Cancer Immunotherapy. Cancer Cell; **33**(4):570-580. [2018]

3. Zhang X, Ning Z, Mayne J, Figeys D. Clinical Microbiome Analysis by Mass Spectrometry-Based Metaproteomics. Annu Rev Anal Chem (Palo Alto Calif); **18**(1):149-172. [2025]

Targeted Gene Correction of the Artemis (DCLRE1C) c.1543G>T Mutation in CD34 Cells Using CRISPR-Cas9 and HDR

Anna Laufer

Advised under Prof. Ayal Hendel, Dr. Michael Rosenberg, PhD student Nimrod Ben-Haim and researcher Esther Wasserstein

Severe combined immunodeficiency (SCID) is a group of rare genetic disorders marked by impaired T and B cell development, leading to life-threatening infections early in life. One form is caused by mutations in the *DCLRE1C* gene, which encodes Artemis, a key enzyme in V(D)J recombination during lymphocyte development. Mutations, such as the c.1543G>T nonsense mutation in the DCLRE1C gene, prevent the formation of a functional adaptive immune system. The only curative treatment, hematopoietic stem cell transplantation (HSCT), is limited by donor availability and risks such as graft-versus-host disease. As an alternative, gene therapy using autologous hematopoietic stem and progenitor cells (HSPCs) offers a safer, more targeted option. CRISPR-Cas9 gene editing, paired with homology-directed repair (HDR), enables precise correction of disease-causing mutations while preserving natural gene regulation, and holds strong potential for developing personalized therapies for SCID.

This project tested the feasibility of correcting the c.1543G>T nonsense mutation in the DCLRE1C gene. Using CRISPR-Cas9-mediated homology-directed repair (HDR), this mutation was targeted in human $CD34^+$ HSPCs. A corrective G base

was introduced at the Cas9 cut site to restore the wild-type sequence (Figure 1). Additionally, two complementary HDR templates were compared, corresponding to the top and bottom DNA strands, to assess strand preference in precise gene correction.

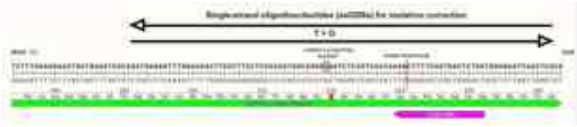


Figure 1. Targeted correction of Artemis (DCLRE1C) c.1543G>T mutation using HDR template

The editing reagents were delivered into CD34^{+} HSPCs using electroporation, a technique that briefly applies an electrical pulse to create temporary pores in the cell membrane, allowing uptake of the Cas9 ribonucleoprotein complexes and HDR templates. Following electroporation, cells were allowed to recover in culture before genomic DNA was extracted for analysis by Sanger sequencing and next-generation sequencing (NGS). Cas9 activity was readily detected, with overlapping peaks in edited samples compared to distinct traces in control samples (Figure 2).

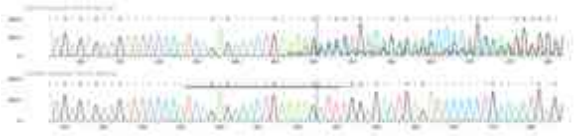


Figure 2. Sanger sequencing of Artemis in edited vs. control CD34^{+} cells

Quantification revealed robust indel formation across all Cas9-treated samples, confirming efficient editing. However, HDR-mediated repair occurred only with the top-strand-oriented template, achieving a correction of ~18–22%. The bottom-strand-oriented template yielded no detectable repair (Figure 3).

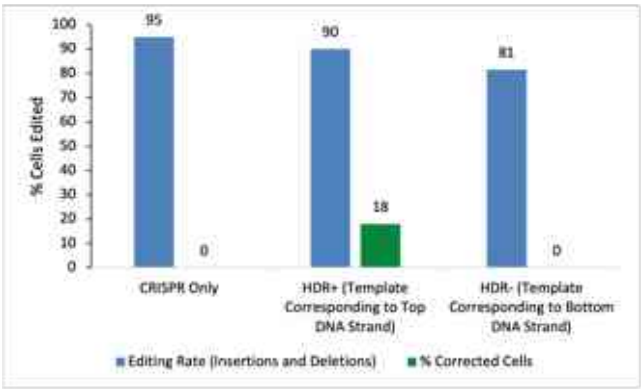


Figure 3. Efficiency of CRISPR-Cas9 and HDR template Eediting in CD34^{+} cells based on Sanger sequencing (n=2).

NGS sequencing provides more precise data. Untreated cells showed 98.5% alignment to the reference sequence (Figure 4A). Cas9-only and HDR^{-} conditions were dominated by non-homologous end joining (NHEJ), with >80% indel frequency and minimal HDR (Figure 4B). By contrast, cells treated with Cas9 and the HDR^{+} template exhibited 22.4% precise correction, with the remaining alleles repaired by NHEJ (Figure 4C).

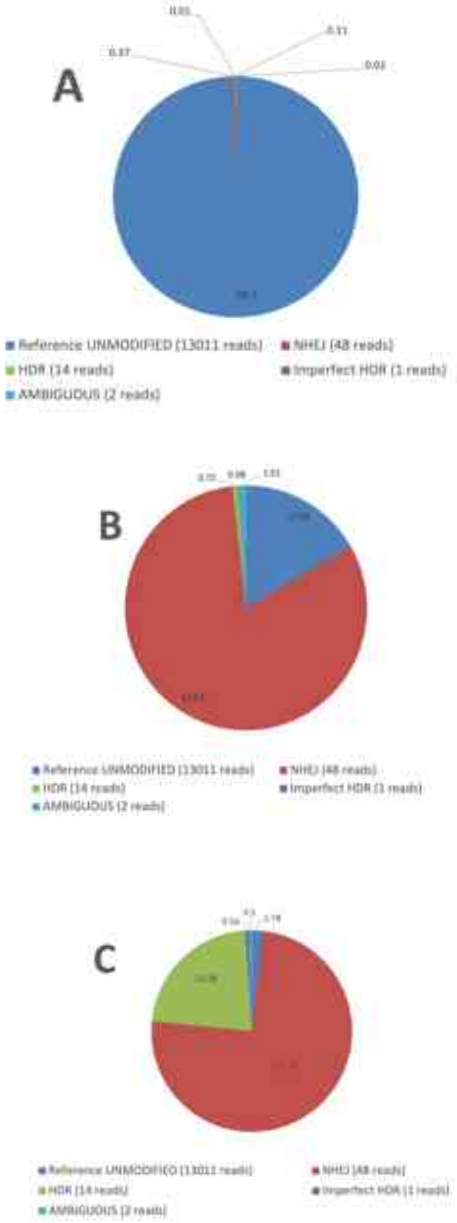


Figure 4. Next-generation sequencing analysis of editing outcomes in CD34⁺ HSPCs (Representative Data).

These findings demonstrate proof-of-concept for correcting Artemis-SCID mutations in clinically relevant stem cell populations. Even modest levels of precise repair are predicted to restore immune function, making this a promising foundation for autologous therapies. The results highlight the importance of template design and underscore the challenges of improving HDR efficiency, minimizing off-target edits, and adapting this approach for clinical application.

Cython Optimization When Creating Gene-Coexpression Matrices

Ariel Melnitsky

Advised under Dr. Binyamin Knisbacher and MSc student Nadav Klein

Quantifying heterogeneity, the diversity of gene expression patterns, is a critical tool in bioinformatics; it provides insights into the mechanisms underlying genetic change. For instance, older cells show elevated heterogeneity, revealing that DNA replication and repair in older cells cause an increased amount of genetic mutations. Since cancerous cells exhibit a similarly impaired capacity, this study utilizes aging as a model system for cancer research. Methodologically, this study develops a framework to build single-cell gene coexpression networks (GCNs) from scRNA-seq data. In these networks, genes are represented as nodes and edges between nodes indicate the likelihood of coexpression. Comparing the divergence of network characteristics from the GCNs of young and old cells provide insights into the

biological systems that drive genetic alterations during aging. Applying this methodology to analyze single-cell coexpression networks made from cancerous cell data is a promising avenue for future research in bioinformatics and cancer biology.

The creation of gene coexpression networks (GCNs) relies on two critical components: a coexpression matrix and a corresponding p-value matrix. The coexpression matrix represents the probability that two genes are coexpressed within the same cell, while the p-value matrix provides the statistical significance of these probabilities. In Python, generating these matrices involves significant computational overhead, slowing down the construction of GCNs. The underperformance of Python is amplified by the large-scale, high-dimensional single-cell datasets. Improving the speed of this process is therefore essential for efficient large-scale analyses.

Cython allows seamless integration of C-level operations with Python and therefore exhibits significant speed optimizations. More specifically, the reduction in runtime was accomplished through the use of 'cdef' static typing to convert Python objects and loops into C-level operations, typed NumPy arrays for fast contiguous memory access, and Cython-level loops to eliminate interpreter overhead. Additional improvements, such as streaming mean and variance calculations to reduce memory usage and the integration of optimized correlation libraries (CorALS), enable rapid computation of correlations in high-dimensional datasets, substantially accelerating the speedy construction of GCNs. The Cython-optimized implementation demonstrated a substantial performance improvement, reducing runtime by 34 seconds compared to the original Python code's 1 minute and 57 seconds. These results represent the average of seven independent runs, ensuring the reliability and accuracy of the measurements. Future research in this area should explore the complete implementation in C, allowing tighter integration with hardware and eliminating all Python bottlenecks.

Modeling CAKUT Pathogenesis Using Kidney Organoids with Heterozygous DACT1 Mutations

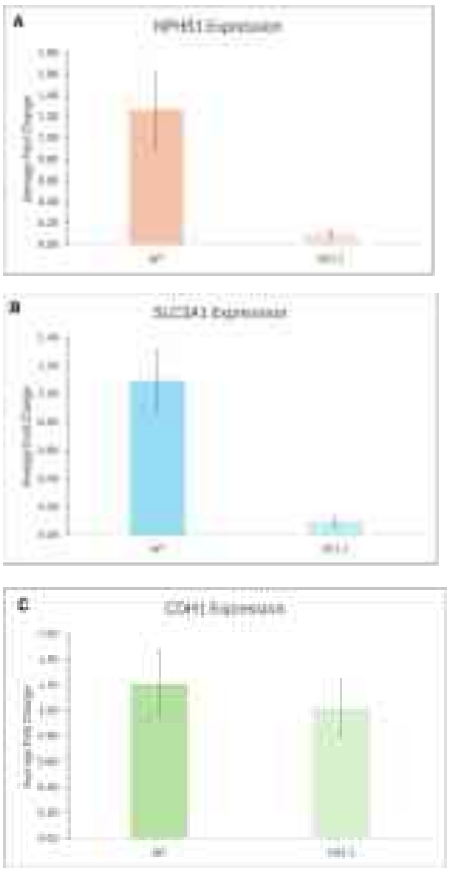
Rachel Schwartz

Advised under Dr. Achia Urbach, Dr. Leah Armon and PhD student Daniella Folkman Genet

During embryonic development, kidneys form through nephrogenesis, a process in which specialized tissues generate filtration units known as nephrons. Disruptions in this process can lead to congenital anomalies of the kidney and urinary tract (CAKUT), a major cause of chronic kidney disease. Although genetic and linked to CAKUT, but the effects of environmental factors contribute to CAKUT, its underlying molecular mechanisms remain unclear. DACT1, a gene involved in tissue organization and cytoskeletal regulation, has been heterozygous mutations on

nephron development remain unclear.

To investigate these effects, we used CRISPR-Cas9 to introduce a patient-specific heterozygous DACT1 mutation into human pluripotent stem cells (hPSCs), which were subsequently differentiated into three-dimensional kidney organoids. RNA was extracted from wild-type (WT) and heterozygous (HET) organoids, and RT-qPCR was used to measure the expression of genes representing key nephron segments: NPHS1 (podocytes), SLC3A1 (proximal tubules), CDH1 (distal tubules), and GATA3 (distal tubules and stromal cells).



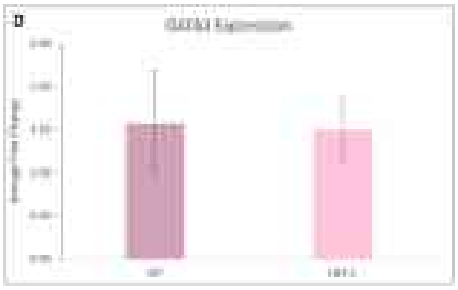


Figure 1. (A-D): Gene expression analysis of nephron segment markers in WT and DACT1 heterozygous kidney organoids

Compared to the wild type, DACT1 heterozygous organoids displayed significantly reduced NPHS1 and SLC3A1 expression, indicating impaired podocyte formation and proximal tubule differentiation. CDH1 expression showed moderate reduction, while GATA3 levels remained unchanged. These findings demonstrate that even a single DACT1 variant can alter nephron-specific gene expression and disrupt kidney development. By modeling patient-derived mutations in kidney organoids, this study provides insights into the molecular basis of CAKUT and provides a valuable platform for uncovering the molecular mechanisms driving congenital kidney defects.

PHYSICS, CHEMISTRY, AND MATHEMATICS



Noah Bodner, Nissim Farhy, Devora Weinstein, Batyah Jasper, Hannah Goykadosh, Elza Koslowe

How Free are the Oceans and Marine Clouds from Continental Influence?

Noah Bodner

Advised under Dr. Tom Goren and Dr. Goutam Choudhury

Earth's inhabitability is a result of a balance of Earth's energy budget: a balance of the radiation it receives and emits. Earth receives constant radiation from the Sun in the shortwave region of the radiation spectrum and emits longwave radiation from its surface. The Sun is the major source of heat, without which humans would not be able to live. Yet, without Earth's atmosphere, much of the heat would escape into space. In this way, the atmosphere serves as Earth's blanket, trapping in heat to keep us warm. But Earth needs a way to moderate how much energy it receives so as not to become too hot. It does so by reflecting incoming shortwave radiation. Earth's reflectivity is referred to as albedo, which is the percentage of solar radiation that is reflected by a surface. Notably, the oceans, which appear relatively dark on satellite imagery and cover more than 70 percent of Earth's surface, have albedo values between 0.05 and 0.10. This means that our oceans reflect between five and ten percent of sunlight that hits them, absorbing an impressive 90 to 95 percent of

incoming radiation. So, any change of albedo over the oceans has a meaningful impact on the overall albedo of the Earth, which affects Earth's energy balance.

Clouds, especially low-level liquid clouds over the oceans can have very high albedo values with respect to the oceans below. In this way, clouds are acting as a mirror for Earth, reflecting much of the incident solar radiation and preventing the absorption of radiation by Earth's surface. And so, the prevalence of marine clouds, their characteristics, and ways in which their properties are altered is important to understanding Earth's energy balance in the short and long term.

One variable that affects the properties and formation of these marine clouds is the extent to which cloud condensation nuclei (CCN) are present over the oceans. CCN are aerosol particles, around which water vapor condenses to form water droplets. These water droplets are what comprise clouds. Having more water droplets in a cloud increases the clouds albedo, and having a higher amount of CCN present in a cloud will result in a higher number of water droplets [1]. CCN have many sources and are in relatively higher concentration over land than over the ocean, where the air is more pristine, as most CCN source regions are land-based.

Because clouds strongly influence Earth's albedo, and because the number of CCN in a cloud can result in the increase in the cloud's albedo, it is essential to investigate the extent to which aerosols from land reach over the ocean.

As a proxy to follow the trajectory of aerosols from land to over the ocean, a lagrangian model tracking the trajectory of continental air parcels, which contain aerosol particles, over a seven day period, was used. The model covers 21 years, spanning from 2004-2024. The output data was processed to include only lower levels of the atmosphere, the region in which CCN interact with marine clouds. To determine the portions over the ocean that are impacted by continental air parcels, the air parcel trajectory counts per given coordinate averaged over the time scale were calculated and plotted. The same was done for mean trajectory counts for each season. Annual and seasonal trends in trajectory counts on the decadal scale were also calculated and plotted. The trends, compared with the trajectory counts, help highlight regions in which the trend is significant.

Preliminary analysis shows vast regions of the ocean with very low trajectory counts within a seven-day period (Figure 1). The regions not reached by aerosols are pristine as after the seven-day period it is assumed that a significant portion of the aerosols within the air parcel are no longer present. Further, the results highlight regions of the oceans with very high counts of trajectories, indicating regions that are frequently impacted by land aerosols. A prime example of this is off the West coast of Africa over the Atlantic Ocean (Figure 2). We then

analyzed the trajectory count trends across the globe. This revealed that in the region off the West coast of Africa there is also an increasing trend of trajectory counts over the 21 years of data (Figure 3). This region is known for the site where significant amounts of Sahara dust is transported westward across the Atlantic Ocean. These regions of high trajectory counts invite further research on its effect on marine clouds and their reflectivity. The regions of high mean trajectory counts also present locations of prevalent wind patterns. The trends plot suggests wind patterns may be shifting which invites further research on how global wind patterns have changed over time, and how climate change relates to these trends. Further research is needed and can include combining cloud fraction data and trajectory count data to highlight regions with significant cooccurance.

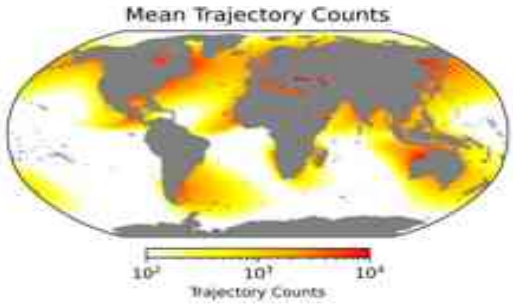


Figure 1. Mean trajectory Counts model output from 2004-2024. Shows extent to which aerosols move over the oceans. Reveals polluted vs. pristine locations.

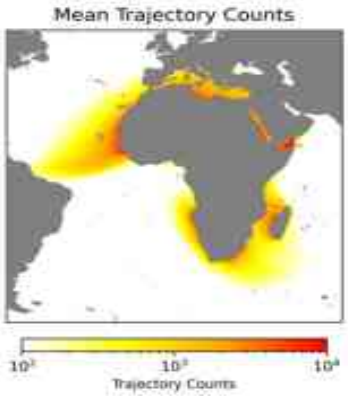


Figure 2. Mean trajectory counts originating from the African Continent, showing air transport across the Atlantic Ocean from the West coast.

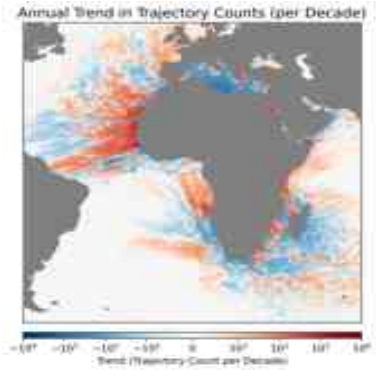


Figure 3. Annual trend in trajectory counts originating from the African Continent. Features an increase in air parcel trajectory counts off the West coast of Africa, over the Atlantic Ocean.

1. S. Twomey, Pollution and the planetary albedo, Atmos. Environ., **8**, 1251–1256 (1974).

Radiative Diffusion Utilitizing a Monte Carlo Simulation

Nissim Farhy
Advised under Prof. Asaf Pe'er

Radiative diffusion, as it relates to stellar and black hole energy transport, is a critical process to understand in astrophysics. Stars,

as well as accretion disks that surround black holes, emit energy which is generated as heat and released as light (radiation). In both cases photons travel in a "random walk" colliding with neighboring particles as they depart. The distance between each collision is called the mean-free-path (λ) which changes as the photon leaves from the stellar interior or the accretion disk.

My research simplifies these complex movements using a Monte Carlo (MC) simulation — a simple statistical analysis tool — to compare the differences in a diffusive and free regime. The simulation uses a center-launched photon model to systematically explore the transition between the ballistic ($\tau \ll 1$) and diffusive ($\tau \gg 1$) regimes by varying λ across four orders of magnitude and analyzes photon flux over the different opacities.

To create the simulation, I used Python to imagine a single photon centrally-launched in a spherical cloud or radius R_0, and explored how the optical depth (τ) affects the photons' escape time and behavior. Using the simple equation $\lambda = R_0/\tau$ I varied the optical depth and calculated the mean number of steps and the mean escape time for 10,000 photons at each optical depth.

As predicted, in the free regime photons escaped at time R_0/c as if there were no other particles. Conversely, in the diffusive regime, photons escape time is longer and proportional to τ. Using mathematical formulas this simplified to $(\tau R_0)/2c$. The transition between regimes occurs sharply (90% accuracy) near $\tau = 1$, validating the expected theoretical relationship$\langle t \rangle \propto \tau$ for the diffusive limit. It is important to note that other simulations I conducted demonstrated that a *randomly* placed photon diffuses to the square of the opacity τ^2 and leaves a constant of ¾.

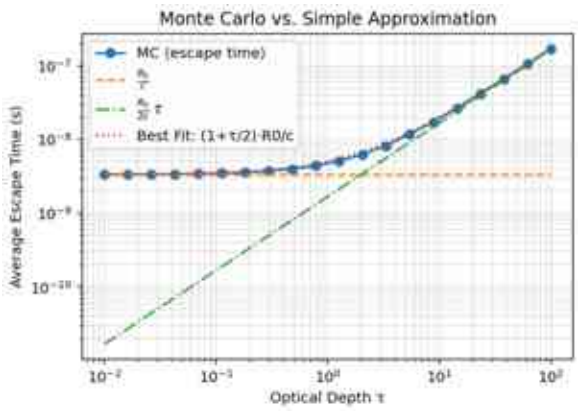


Figure 1. The primary Monte Carlo simulation vs. approximate equation

Further analysis found that the photons escaping per unit area, also known as the photon flux, demonstrates that the energy escape rate declines sharply, scaling inversely with the mean escape time Flux $\propto 1/\langle \tau \rangle$.

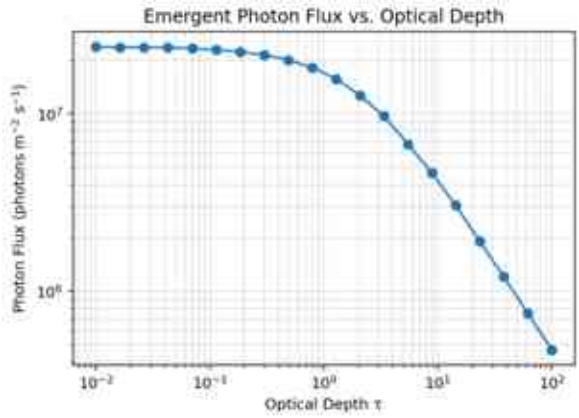


Figure 2. Emergent photon flux vs. optical depth

Finally I plotted the cumulative distribution function (CDF), plotting the percentage of photons that escape over time. These histograms prove that photons in radiative diffusion take on the bell-esque diffusion of a gamma distribution, demonstrating that photons exhibit a true statistical random walk. These analyses provide the framework to understand energy transfer and radiative diffusion in the surrounding of black holes and stars.

Over and Over Again: Exploring Iterations of Voronoi Diagrams

Hannah Goykadosh
Advised under Professor Emanuel (Menachem) Lazar

Voronoi diagrams divide space into regions around a set of given objects. Each region contains all points in the space which are closer to that object than to any other object. One example of where a Voronoi diagram could be useful is when trying to figure out where everyone in a neighborhood should send their mail. Imagine this neighborhood has three post offices called A, B, and C. All the space in the neighborhood closer to post office A compared to B or C would be in the Voronoi cell surrounding post office A and the resident's mail would be sent there. An example of a Voronoi Diagram is shown below in Figure 1, each yellow dot represents an object, and each blue polygon is the Voronoi cell about that object.

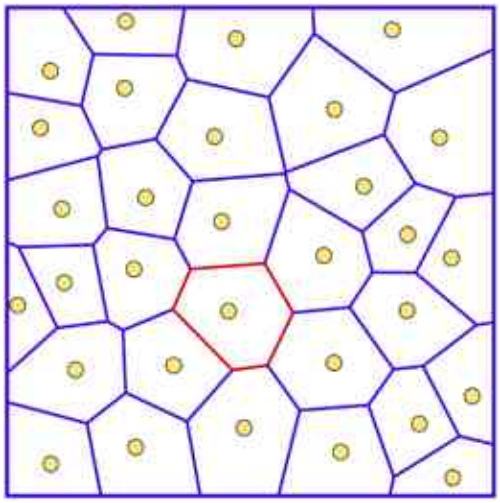


Figure 1. Voronoi Diagram [1].

Voronoi diagrams are incredibly useful tools for analyzing many different types of physical systems in which each point can represent an object such as an atom. These diagrams allow for the analysis of large data sets to be boiled down to questions about polygons and polyhedra and provide insight into various physical systems [1].

The objective of my research was to investigate the behavior of a Voronoi cell under an iterative reconstruction. To build this iterative construction, the cell surrounding a central point (*C*) was isolated and the vertices of this cell were used as new object points to generate the successive Voronoi Diagram. To investigate this iteration, visual simulations of this construction were observed.

I first made observations in the two-dimensional case through visual simulation. Every Voronoi cell is a convex *n*-gon and I observed that each subsequent iteration remained a *n*-gon with the same *n* number of sides and vertices as the original

cell. Notably, each *n* iteration of the mapping will result in a similar *n*-gon rotated and scaled by uniform factors. An example of this iterative mapping of the Voronoi cell can be seen below in Figure 2. The red central point represents locus point *C.* This example begins with the original cell that is a triangle, each successive iteration remains a triangle. The fourth iteration is similar to the first iteration rotated and scaled about the central point. In this figure iterations are normalized for the sake of visual clarity.

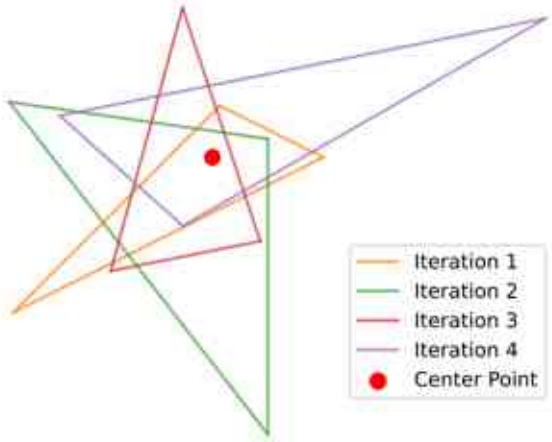


Figure 2. Iterative mapping of the Voronoi cell

This periodic behavior can be proven by understanding the Voronoi diagram as a circumcenter map. A circumcircle is the unique circle that passes through any three given points. The circumcenter map is a prescription for creating a new polygon by taking the triangulation of an arbitrary locus point and the two adjacent points in the original polygon, drawing the circumcircles of each of these triangulations and connecting the centers of these circumcircles in order. This construction is called the pedal polygon. The pedal polygon drawn by the circumcircle map is the exact same polygon described by the Voronoi iterative map. It is a special case of this map where the locus point is the central point *C* the Voronoi map is iterating over.

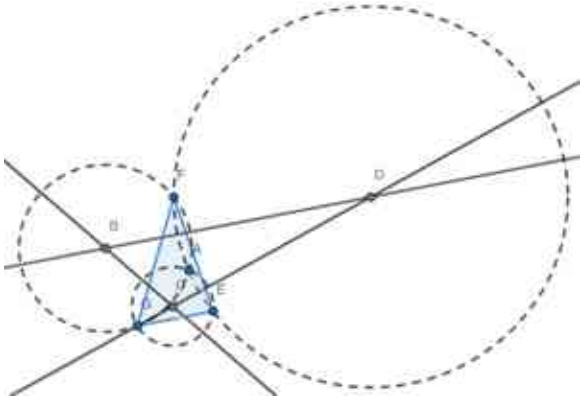


Figure 3. Example of Circumcircle map. The Circumcircle map is generated over triangle EFG about locus A, resulting in a new triangle BCD, which is also the new Voronoi cell about point A, in consideration to points E,F,G.

This was proven by B.M. Madison in 1940 through a series of geometric constructions [2]. There are several intriguing cases and observations regarding this map and its converging and diverging regions as well as special cases of periodicity that are explored and proven in [3].

I expanded my research of this iterative Voronoi map to the third dimension. In two dimensions the proven patterns are clear and are generalized to all polygons. However, in the third dimension these generalizations do not hold true. Excluding some interesting cases that will be

discussed, generally iterating over a three-dimensional polyhedron cell about *C* results in a polyhedron with an increasing number of sides, edges, and vertices in each iteration or the points become coplanar and the polyhedron degenerates.

An initial tetrahedron configuration may either degenerate or remain a tetrahedron under iteration. There is no other possibility for a polyhedron under this iteration. An equilateral tetrahedron transforms into a similar tetrahedron rotating about the centroid. This rotation results in two possible locations for the equilateral tetrahedron.

An initial configuration of a cell that is a perfect cube results in two possible polyhedrons, an eight or twelve faced polyhedron. The iteration alternates between these two possibilities. In the cell that begins with a five faced triangular prism, the map iterates between a six or eight faced polyhedron. This is not assumed for all possible polyhedron initial configurations; these are unique cases.

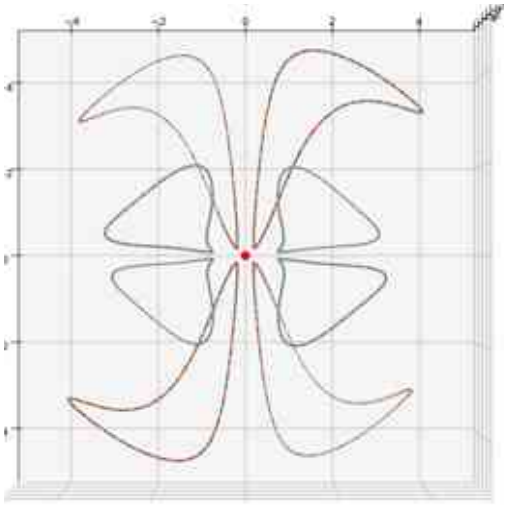


Figure 4. Plotting of the Curves Generated by the Three-Dimensional Iterative Mapping of the Voronoi Cell with the Initial Configuration of a Tetrahedron.

The most fascinating results arise from the configuration of the initial points as a tetrahedron. As discussed earlier, the successive cells remain a tetrahedron. When plotted, the vertices of each of these mapping form eight distinct discrete curves upon which the vertices of all iterations fall. The curves can be categorized into two distinct groups alternating on iterations. Each of these two groups of curves contain are self-similar. While these curves are not yet well understood, the existence is clear in the visual simulation results and are incredibly intriguing and show a clear underlying periodic pattern that lays a basis for future research.

This project provided valuable insights into the geometry of iterated Voronoi diagrams. Future work will focus on rigorously proving these results in three dimensions, exploring more complex polyhedra, and characterizing the curves generated by the tetrahedral iteration.

1. E. Lazar, J. Lu, C. Rycroft. Am. J. Phys*.,* **90**, 469–480 (2022).

2. B. M. Madison. The American Mathematical Monthly, **47**, 462-466 (1940).

3. N. McDonald, R. Garcia, D. Reznik. The Mathematical Intelligencer, **45**, 232–241 (2023).

Performance Evaluation of Platinum Catalysts in PEM Hydrogen Fuel Cells

Batyah Jasper
Advised under Prof. Lior Elbaz and M.Sc. student Eliana Lebowitz

The development of renewable and widely accessible energy systems is limited by the challenge of storage. Energy that is dispersed through the grid must be utilized immediately upon generation and requires expensive infrastructure to reach all areas where it is needed. The hydrogen economy has emerged as a method of capturing excess energy by using it to power endothermic electrolysis, which splits water into hydrogen and oxygen. The hydrogen gas produced serves as the fuel for hydrogen fuel cells, which offer a more sustainable and energy-dense method of storing and transporting energy than traditional batteries. Once the hydrogen enters the proton exchange membrane (PEM) cell, it is split into a proton and electron and later recombined with oxygen, releasing energy in the process and water as the only byproduct. Large-scale implementation is currently hindered by the high cost of platinum catalysts necessary for the oxygen reduction reaction (ORR) at the cathode.

My research investigated the performance of platinum catalysts to inform the design characteristics required for alternative catalyst materials, such as aerogels, to be commercially viable. Platinum nanoparticles were deposited on a carbon black (XC72) backbone using the incipient wetness impregnation technique. The highly porous nature of the carbon backbone provides the critical active binding sites for the ORR to occur. Energy dispersive x-ray spectroscopy (EDAX) verified the proper elemental composition of the catalyst, being 76.3% carbon and 20.3% Pt with minimal contaminants. The porous structure, crystallite size, and facet orientation were characterized using x-ray diffraction (XRD) and a scanning electron microscope, shown in Figures 1 and 2.

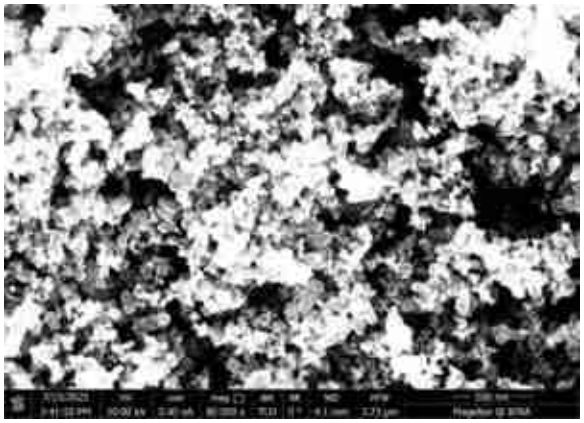


Figure 1: Scanning electron microscope image of highly porous catalyst on 500 nanometer scale

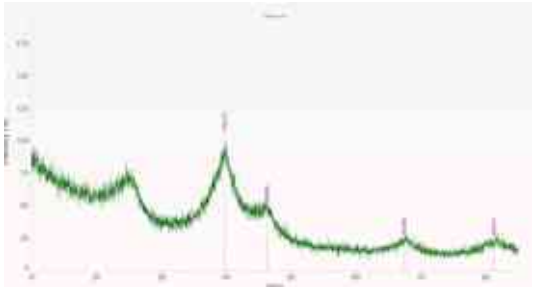


Figure 2: X-ray diffraction spectrum of crystallite surface facets

XRD confirmed a dominant (111) surface facet orientation, which is the most catalytically favorable configuration for ORR, and the Scherrer equation indicated an average crystallite size of 1.15 nm.

A three electrode half cell was utilized to independently evaluate the performance of the catalyst during the ORR through cyclic voltammetry (CV) and rotating disk electrode (RDE) measurements. CV measurements of the hydrogen desorption,

which is directly proportional to the quantity of active binding sites, allowed for the determination of the electrochemical surface area (ECSA). The measured ECSA was 67.2 cm^2, more than two orders of magnitude greater than the geometric surface area of 0.25 cm^2. This dramatic increase indicates that the catalyst's highly porous structure exposed vastly more active Pt surface sites than suggested by its physical footprint, thereby enhancing ORR performance. The current generated from input voltage at various speeds (rpm) during the RDE experiment is depicted in Figure 3.

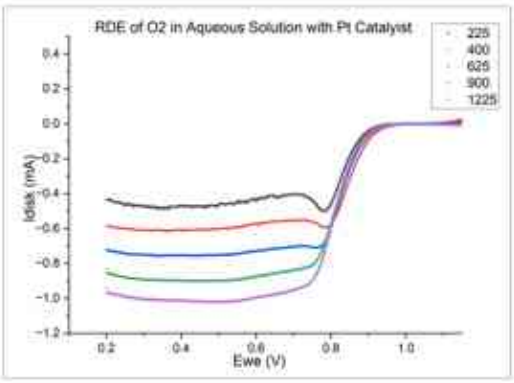


Figure 3: RDE of oxygen in aqueous solution in three electrode half fuel cell at varying speeds

As expected, the limiting current increased with electrode rotation as forced convection minimized the diffusion layer. Levich's equation demonstrated that the ORR proceeded primarily through the desirable four-electron pathway to water, which is twice as efficient as the undesirable side reaction forming hydrogen peroxide, with an average value of 3.7 electrons transferred across all speeds. Finally, the catalyst was tested in a complete PEM fuel cell, demonstrating stable and effective catalytic performance under realistic operating conditions and resistances, validating the relevance of the half cell findings. Ultimately, these findings provide insight into the key properties of catalyst structure, size, crystallite facet orientation, and electrochemical performance that must be replicated or improved in next-generation non-platinum catalysts for fuel cell applications.

Implementing Secure Multiparty Protocols in Python

Elza Koslowe
Advised under Professor Gilad Asharov

Cryptography is the area of research dealing with secure communication of private information in the presence of an adversary. In secure computation, parties wish to compute a joint function of their inputs while keeping their inputs private. For instance, hospitals can collaboratively analyze patient data to detect disease trends without disclosing individual medical records, or competing companies can compute the average salary across firms without exposing specific salaries.

My research was focused on a specific cryptographic protocol designed by a number of cryptography researchers, including Professor Ariel Nof of Bar-Ilan University, who helped advise me on my project. This protocol allows for secure computation between three parties, assuming that one of them might be corrupt.

The objective of my research was to implement functions of this protocol using the Python programming language. The

initial steps of the research involved understanding the math behind the protocol, as well as understanding the Advanced Encryption System (AES) and how to use it in my code. Once I understood this, I was able to create a Python class, such that each party of the three-party share is an object of the class, and is able to interact with the other parties using methods of the class. These methods include addition and multiplication, which are the building blocks of any secure computation; once addition and multiplication are possible, any secure computation can be achieved.

Phosphorylation on the Molecular Level in E2F8 Using NMR Spectroscopy

Devora Weinstein

Advised under Prof. Jordan Chill and Dr. Inbal Sher

Background

E2F8 [1] is a member of the E2F family of transcription factors, which regulate many essential cell functions. The family divides into canonical E2Fs (E2F1–6) and atypical E2Fs (E2F7–8), with E2F8 showing strong structural and functional similarity to E2F7. qPCR studies reveal high expression of E2F7/8 in late S phase, followed by degradation in G2 through phosphorylation and ubiquitination. Phosphorylation, the addition of a phosphate group to serine, threonine, or tyrosine residues, is a key regulatory mechanism. E2F8 contains several phosphorylation sites, including T20 and T44, with evidence suggesting T20 phosphorylation depends on T44. To investigate this, recombinant E2F8 proteins were expressed in *E. coli*. Because phosphorylation is a eukaryotic process, site-directed mutations were introduced to mimic it: T-to-A creates a “phosphodead” residue, while T-to-D introduces a negative , also known as “phosphomimicking”.

Methods

The E2F8 gene was cloned into an expression vector, transformed into E. coli, and expressed in isotopically labeled media. Purification was performed using affinity chromatography, followed by size-exclusion chromatography and dialysis. Protein quality was assessed by SDS-PAGE. Nuclear magnetic resonance (NMR) spectroscopy was used to obtain $^{1}H–^{15}N$ HSQC “fingerprint” spectra of wild-type and mutant proteins (Figure 1). A kinase assay was applied to phosphorylate the wild-type protein, and structural changes were monitored over time.

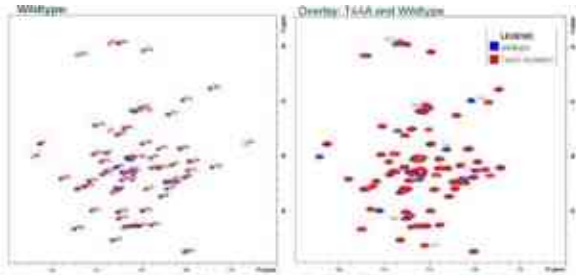


Figure 1. Fingerprint spectra of E2F8 wildtype and T44A mutation

Results

The fingerprint spectra of T44A compared with wild-type revealed significant chemical shift perturbations in residues surrounding the mutation site. Amino acids with a high CSP (chemical shift perturbation) are highlighted with arrows in the overlay.

Residues that are in the surrounding of the mutated amino acid were changed as a result of the mutation. There are also some changes that can be observed in G16/L17, which can suggest that these regions are in proximity to each other. These spectra are the bases for the phosphorylation experiments.

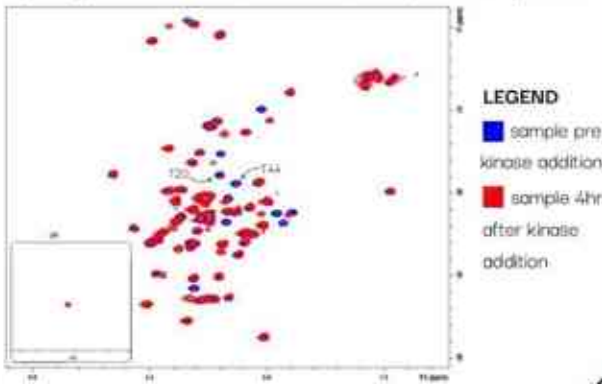


Figure 2. NMR spectrum of E2F8 wildtype before addition of kinase and after 4 hours

To investigate how phosphorylation affects protein structure, we added a protein kinase and monitored changes over time using NMR spectroscopy. The spectrum in Figure 2 shows an overlay of the sample before kinase addition and after 4 hours. Peaks corresponding to amino acids T44 and T20 are highlighted. Both signals decrease significantly in intensity after 4 hours, and a new peak appears elsewhere in the spectrum. Further analysis is required to confirm the identity of the new peak, still these preliminary results indicate that phosphorylation induces structural changes in the protein.

Conclusion

Overall, these experiments help shed light on the role of E2F8 in the cell cycle. The phosphorylation experiments will help determine if phosphorylation in the protein is residue dependent. In the future, we will repeat the phosphorylation experiment with T44D and T44A, to see how/if phosphorylation occurs. These results are also used to compare the phosphorylation trends of E2F7 and E2F8. These proteins are thought to be derived from a common ancestor, and they work synergistically in the cell; however, they do not play the exact same role in the cell. Uncovering their differing phosphorylation patterns (timing, etc.) may shed light into more deeply understanding their respective roles in the cell cycle.

1. Lv, Yi, et al. "E2F8 is a potential therapeutic target for hepatocellular carcinoma." Journal of Cancer, **8** (7), 1205-1213 (2017).