Bar-Ilan–Yeshiva University

Summer Science Research Internship Program 2022

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, Joan Seidel, Zoltan Erenyi Fund and Yeshiva University, students gain invaluable laboratory skills, along with an unforgettable summer experience.

Program Director: Prof. Arlene Wilson-Gordon and Prof. Ari Zivotofsky

Av and Em Bayit: Rav Shlomo and Adi Anapolle

TABLE OF CONTENTS

Editor: Abby Goldberg
Contributing Editors: Avigayil Roffe and Noa Atar

BRAIN SCIENCES

(L-R) Abby Goldberg, Shoshana Linfield, Sharon Khalil, Yaelle Akhavan, Katherine Belilty, Noa Atar, Noam Putterman, Eitan Maron

Studies Related to Emotional Processing Conducted in a Psychopathological Laboratory

Yaelle Akhavan, Noa R. Atar, Sharon Khalil

Advised under Prof. Eva Gilboa-Schechtman, PhD student Dan E. Hay, and Master's students Noa Perets, Raquel Landau, and Itamar Zalkind

Emotional Congruence and Depression

Depression is one of the most prevalent and severe mental disorders today. Due to its prevalence, there is a lack of resources and therapists to help treat all patients. Therefore, the aim of this study is to create an alternative intake resource to be able to treat patients on a need basis. AI is a new possible alternative that can be used to solve this dilemma. Because depression is accompanied by many nonverbal cues, this study works with a *Facial Expression Recognition Software* (FaceReader 9) to determine whether it could be used as a tool in helping diagnose depression [1]. This study runs

short intake videos (around 25 minutes) of clinical patients through *FaceReader*, which analyzes the six basic universal facial expressions in order to detect their valence (e.g., happy vs. sad) and arousal (intensity of emotion). The valence and arousal are then compared to the content of speech to determine the level of emotional congruence.

When a healthy individual speaks, the assumption is that the emotion conveyed in the facial expressions are congruent with the emotion conveyed by the content of speech, known as emotional congruence. When the facial expressions do not match the content of speech, there is said to be emotional incongruence, which is associated with an impairment in the individual's psychological functioning, and potentially be indicative of worse mental health [1].

The authors hypothesized that the level of congruence between the different emotional channels (facial expression and content of speech) can contribute to the clinical assessment of depression, even more than the BDI and Hamilton Assessment do; the higher the incongruence detected by the analysis of *FaceReader* and speech content, the more severe depression an individual may experience. Additionally, the initial level of congruence before therapy begins is expected to predict therapy outcome, as well as symptom severity after the therapy ends.

This study is the first step towards determining whether *Facereader* can be a useful tool in helping diagnose depression more efficiently and promptly, considering the difficulty in getting in touch with a mental health professional in a timely manner.

Figure 1. Screenshot of FaceReader detecting the valence and arousal.

Meta Analysis: Effects of Exclusion on Individuals with BPD

Interpersonal connections are essential for the emotional and psychological wellbeing of all individuals. From an evolutionary standpoint, social rejection is seen as a threat to human survival, causing psychological distress and even pain. In 2009, Williams suggested that the immediate impact of exclusion (*Reflexive Stage*) is universal and has no variability among individuals (healthy control and psychopathological states alike) [2]. Reinhard, however, suggested that different psychopathological disorders can, in fact, enhance the effects of exclusion. Reinhard, describes this interaction between exclusion and psychopathological disorders as a "vicious circle," and proposes that different disorders will lead to different, and possibly more severe, reactions, as well as less effective coping strategies in response to the exclusion [3]. This more extreme reaction among clinical patients will compromise their ability to socially interact, leading to further exclusion, hence the name–vicious *circle*. This gap in the research, highlighted by the contrasting views of Williams and Reinhard, emphasizes the importance for further research regarding this topic.

We conducted a meta-analysis which compiled all the results (n=how many studies) from

Cyberball research done with individuals with BPD in order to ascertain how much of an effect exclusion has on those individuals as opposed to healthy individuals. Cyberball is a virtual ball tossing game used to manipulate both social inclusion and exclusion, in a lab setting. Participants are led to believe that they are playing with two or more other players, when in fact, they are computer-generated confederates. People diagnosed with BPD are hypersensitive towards rejection, and therefore it is highly relevant to study how these individuals react to exclusion compared to healthy control.

By gaining a better understanding on how psychopathological individuals respond to exclusion there could be clinical implications that will allow us to better characterize the disorder and the impact of exclusion.

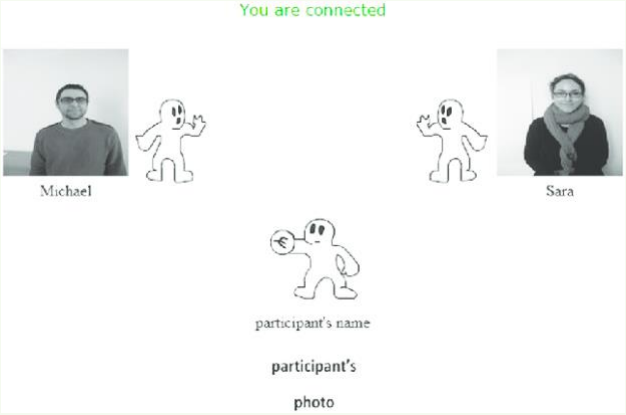


Figure 2. Cyberball screenshot of the three players (one participant; two confederates)

[1] Pampouchidou, A., *et al.* (2017). Automatic assessment of depression based on visual cues: A systematic review. *IEEE Transactions on Affective Computing*, **10**(4), 445-470.

[2] Williams, K. D. (2009). Ostracism: A temporal need-threat model. A*dvances in experimental social psychology,* **41**, 275-314.

[3] Reinhard, M. A., *et al.* (2020). The vicious circle of social exclusion and psychopathology: a systematic review of experimental ostracism research in psychiatric disorders. *European archives of psychiatry and clinical neuroscience,* **270**, 521-532.

The Effect of Buccal Ganglion Stimulation on Buccal Motor Programs in Aplysia

Katherine Belilty

Advised by Prof. Abraham Susswein and Prof. Itay Hurvitz

Feeding patterns in the marine gastropod mollusk Aplysia are demonstrative of broader patterns in consummatory motivational behaviors. Studying these patterns can facilitate a more thorough understanding of the integration of external stimuli into fixed motor patterns and behaviors.

Aplysia are model organisms in neuroscience research because of their simple nervous systems; the connections between various ganglia and their respective behavioral and motor functions have been thoroughly researched. Specifically, the relationship between components relevant to aplysia feeding – the buccal ganglia (BG), cerebral ganglia (CG), and muscles of the buccal mass – has been well established. Aplysia feeding is a simple model of motivational behaviors in animals. Motivational behaviors are those in which an animal takes part in order to maintain homeostasis. These include behaviors to maintain temperature or energy, feeding, bowel movements, mating, etc., and occur in response to a biological need. These behaviors are split into two phases: appetitive and consummatory. The appetitive phase consists of any of a wide variety of movements in search of the stimuli necessary to meet the

determined need and can be interfered with. The subsequent consummatory phase, in contrast, occurs upon detection of the desired stimulus and is very focused, difficult to distract from, and typically consists of repetitive motor patterns. This includes actions such as chewing, cutting, or swallowing. Despite the fact that these consummatory behaviors are repetitive and fixed, sensory information is nonetheless integrated into them. Our research seeks to understand how sensory information is integrated into fixed motor patterns and behaviors.

To study this question, we administered electrical impulses to the buccal ganglia and took electrophysiological measurements from the buccal nerves to analyze the effect on buccal motor programs.

Aplysia feeding begins with the protraction of the jaw and the radula —the toothed structure used by mollusks to grasp food. The radula then attempts to clasp the food, followed by a retraction in which the radula either drives the food further into the esophagus or ejects the food. On a neural level, these feeding patterns are termed buccal motor programs (BMPs). These can be mapped onto the I2 nerve, radula nerve (RN), buccal nerve 1 (BN1), and buccal nerve 2 (BN2). The I2 nerve, RN, and BN1 are active during protraction, whereas the RN and BN2 are active during retraction — I2 activity is indicative of muscle movement during protraction, and the RN displays stronger firing during retraction that correlates to its forceful driving of food into the esophagus. Analyzing the activity of these nerves in conjunction allows us to understand the cycle of feeding through the electrophysiological output. Motor programs are driven by acetylcholine in the cerebral ganglia, which ultimately communicate with the buccal ganglia to generate feeding patterns in the buccal mass of the aplysia. Fictive feeding (correlates of BMPs) can be induced in these nerves by applying the chemical carbachol, a drug that mimics the effects of acetylcholine, onto the cerebral ganglia. By applying carbachol to the cerebral ganglia and taking measurements from the I2 nerve, RN, BN1, and BN2, we were able to study the effects of electrical stimulation on BMPs.

We injected an aplysia with $MgCl_2$ in order to numb the nervous system, dissected it, and removed the buccal and cerebral ganglia. We placed the ganglia on a petri dish and separated each pair of ganglia into distinct chambers so that they were connected but in distinct solutions. We placed electrodes into the I2 nerve, the RN, BN1, and BN2, and measured the output on AxoGraph software. We bathed both pairs of ganglia in seawater and took a control measurement for 10 minutes (Fig. 1).

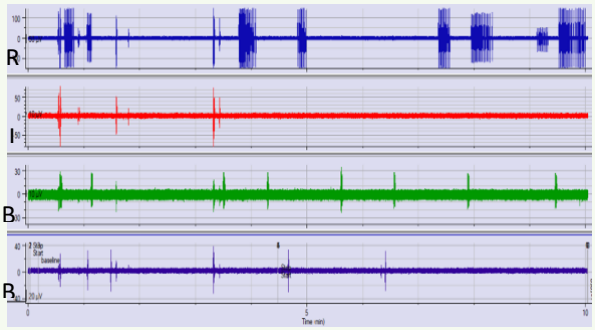


Figure 1: Seawater control measurement from RN, I2, BN1, and BN2 for 10 minutes.

Subsequently, we drained the seawater from the cerebral ganglia, added 1 cc of 1.8×10^{-2} M carbachol to induce BMPs, and took another measurement for 10 minutes (Fig. 2).

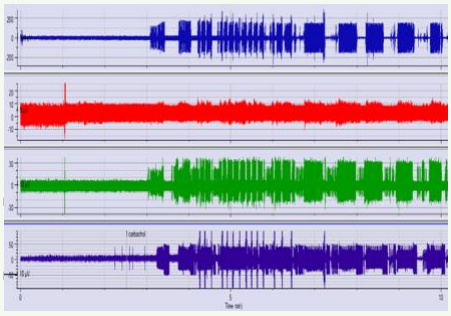


Figure 2: Measurement from RN, I2, BN1, and BN2 for 10 minutes, with carbachol application to CG. After approximately three minutes, the carbachol begins to induce BMPs. Protraction is characterized by RN and BN1 activity, as well as slight I2 activity, for several seconds. This is immediately followed by retraction, characterized by high voltage BN2 units, silence in BN1, and high-voltage, high-speed units in RN.

After this run of carbachol, we washed the CG with seawater and ran another session of carbachol with applied electrical stimulation at several intervals. We applied an electrode to the BG and stimulated the BG after BMPs had been induced (Fig. 3).

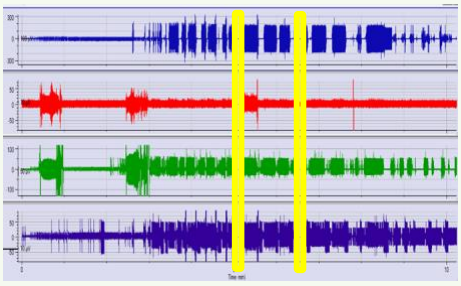


Figure 3: Measurement from RN, I2, BN1, and BN2 for 10 minutes, with carbachol applied to CG and two episodes of electrical stimulation.

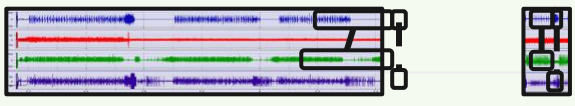


Figure 4: BMP following electrical stimulation (left, 72 second clip) compared to BMP without electrical stimulation (right, 10 second clip). 'P' indicates protraction and 'R' indicates retraction.

Following electrical stimulation, protraction is lengthened significantly; protraction is approximately 18 seconds compared to approximately 5 seconds before electrical stimulation. Furthermore, radula activity is weakened significantly during retraction as is made evident by the lower output voltage, indicating that the radula fails to close during retraction (Fig. 4).

These findings indicate that sensory stimulation of this point in the BG has an effect on feeding behavior and specifically on fixed motor patterns. Subsequent research could include tracing the specific cluster of neurons stimulated in this experiment, analyzing differences between the effects of stimulation on different clusters of neurons in the BG, and ultimately tracing neurotransmitters and molecular correlates to understand the mechanisms by which sensory information is integrated with fixed motor patterns.

The Use of Magnetoencephalography in Brain Studies

Abby Goldberg

Advised under Professor Avi Goldstein and PhD student Barak Atia

MEG, or Magnetoencephalography, is a non-invasive neuroimaging technique. This technique differs from EEG, or electroencephalography, in terms of the data collection. EEG measures the brain's electrical activity, while MEG measures the magnetic oscillations from the brain. This allows for extremely accurate detection of brain activity in terms of both temporal resolution

(milliseconds) as well as spatial resolution (millimeters). The activity of the brain is acquired utilizing a whole-head, 248-channel magnetometer array (4-D Neuroimaging) in a magnetically shielded room. This is done in order to avoid the collection of conflicting data from those electrical fields that can be found in the atmosphere. Since the MEG machine does not produce an anatomic image of the brain, a specific anatomical MRI image of each subject is needed to analyze the source of activity in the brain. In an ideal study, each and every subject would have their own unique MRI scan in order to precisely pinpoint the source of brain activity. Since most studies cannot afford to conduct an MRI scan for each subject, a sample MRI template is used instead. The MRI template needs to be resized and readjusted in order to fit each specific subject's head. As a result of this, the shape and size of every subject's head is measured and recorded at the beginning of most experiments involving MEG. The entire analysis of the data was performed using MEG and EEG data analysis with MNE-Python (Alexandre Gramfort et al., 2013).

The process of converting the raw data from MEG to localizing the source of the brain activity is a multistep process. The first step is recruiting suitable subjects for the experiment. When using MEG technology, the subjects cannot have anything metal in their mouths, since the metal in the mouth can interfere with the data collection of the electrical fields in the brain. Once the subjects are adequately recruited, an electronic digitization system draws a specific head shape file for each subject. After this, the electrical activity of the subjects' brains are analyzed (usually for a one hour period of time or shorter), and the data is collected. Once the data is collected accurately, the process of analyzing the data can begin. The first step of analyzing the data is cleaning the data from any noise that could have interfered with the data. This includes clearing any conflicting electrical signals from the building in which the MEG machine is located, movement from body parts, and any electrical wires or metal. After the noise is cleaned from the data, the data can be processed further. This processing includes firstly fitting each individual head shape file to the template MRI, then loading the data onto that specific head shape file, and then, coregistering the data from the sensor location to the anatomical data that can be used for source localization. Once all of this is done, source localization can be done and conclusions can be drawn based on those calculations.

During my internship this summer at Professor Goldstein's MEG lab at Bar-Ilan University, I aided in processing data and matching the MRI template file to each specific subject's head shape file. The experiment that I processed the data for was Barak Atia's experiment, which researched the effect of positive or negative priming on the brain activity of the followers of charismatic speakers. I helped in the resizing and realigning process to create a unique file that is tailor fit for each subject, using the MNE-Python Analysis toolbox. Additionally, it was the first time in Professor Goldstein's lab that the MNE-Python software was used in tandem with processing the data and matching the head shape files to the MRI template.

MRI template head shape realignment using MNE Python

Maintenance of Sexual Desire in Long Term Couples

Shoshana Linfield

Advised under Prof. Eshkol Rafaeli and PhD student Shira Mond Beker

When approaching the area of relationship research, specific factors must be evaluated within long term couples. For long term couples the sexual relationship is essential for personal well-being as well as overall satisfaction within the relationship. Despite the need for desire within relationships, little is known regarding the factors which help maintain desire within long term relationships. The current project examines the role of certain manifestations of separateness and interdependence between romantic couples in the maintenance of sexual desire in romantic relationships.

Using the Self-Determination Theory (SDT) model, couples are evaluated in three domains- autonomy, competence, and relatedness. Autonomy is the ability to maintain individuality and have self-governance. Competence is the ability to feel capable to overcome challenges when they arise. Relatedness is the ability to feel loved and cared about. Individuals fulfill these needs to varying degrees for their partners. The associations between fulfillment of each SDT factor and relationship satisfaction may differ, and still need to be evaluated.

Another variable which may encompass an experience of dependence and interdependence is a sense of relational power. Power is the partner who possess the greatest amount of dominance within the relationship.

Both the SDT needs as well as relational power will be analyzed alongside intimacy and sexual desire. Intimacy is closeness between couples, both verbal and non-verbal, which does not necessarily have a sensual nature. In contrast, sexual desire is inherently intertwined with sensuality.

Using the Self Determination and Power theories, Professor Rafaeli evaluated long term romantic partners. Specifically, we sought to see if these domains correlate with intimacy and sexual desire in relationships. To evaluate these factors, we coded videos of long-term heterosexual romantic partners. In the videos, couples discussed how they've changed throughout the relationship as well as a conflict in their relationship. Through both verbal and non-verbal communication, the couples' dynamics were noted.

The coding scheme consisted of various variables including SDT needs, power, intimacy, and sexual desire. The SDT variables were defined in the context of fulfilling their partner's needs. Autonomy is supporting

wishes and attitudes of their partner and allowing them to speak freely. Overall, the partner supports the various aspects of their partner's individuality. Competence is the ability to support their partner through listening and offering examples of when their partner has succeeded in life. Relatedness is the partner referring to themselves in terms of couple, i.e., saying 'we', as well other signs of closeness such as smiling or physical touch. Power was evaluated in the context of which partner appears to win arguments and their opinions dominate the conversation. Intimacy is displaying signs of vulnerability by sharing personal feelings as well as physical closeness. Sexual desire manifests through touch in a sensual matter as well as flirtatiousness.

The correlational data will reveal which SDT needs are linked to maintaining sexual desire within long term romantic couples. These results can impact long term sexual relationship counseling in the future by providing empirical data for aiding couples repair relationship by identifying which needs are not sufficiently fulfilled.

B Cells in Alzheimer's Disease

Noam Putterman

Advised under Professor Eitan Okun and Masters Student Khalil Srouji

Alzheimer's disease (AD) [1] is a neurodegenerative disorder which mostly affects elderly individuals. The disease is associated with failure of the immune system to properly remove protein aggregates from brain tissue. These aggregates ("plaques") are composed of amyloid-β (Aβ), which are improperly cleaved peptides originating from the amyloid precursor protein understood, partially due to conflicting scientific evidence. While some studies show that B cells may produce antibodies that reduce Aβ and may also secrete anti-inflammatory cytokines, suggesting that B cells may have an ameliorative or beneficial role, a recent study suggested that B cells may actually be pathogenic in AD. Kim *et al*. [2] utilized multiple models of AD to determine the role of B cells in disease pathogenesis. When a B-cell knockout mutation was introduced into both 3xTgAD and APP/PS1 mice, strains transgenic for AD-associated mutations which exhibit early-onset AD, the absence of B-cells ameliorated the AD phenotype in cognitive (the Morris Water Maze, MWM) and activity-measuring (the Open Field Arena, OFA) tests relative to 3xTgAD and APP/PS1 with preserved B cell numbers. Moreover, 3xTgAD mice treated with anti-B cell antibodies also showed a trend toward increased activity in the OFA, and the amount of Aβ plaques was significantly reduced. These results were validated using mice expressing a more aggressive AD mutation, 5xFAD mice. In this transgenic strain as well, anti-B cell antibodies attenuated the AD phenotype relative to a 5xFAD control and decreased Aβ plaque formation.

Taken together, these results from Kim *et al.* [2] suggest a somewhat counterintuitive conclusion; that is, that B cells may actually *exacerbate* AD progression instead of *retarding* it, as B cells tend to do in other (e.g. infectious) diseases. To further study the roles of B cells in AD, we attempted to accurately profile B cell subsets in various tissues in several models of AD.

There are multiple subtypes of B cells. Based on their cellular markers, B cells largely can be categorized as B1 cells or B2 cells. B1 cells can further be distinguished into B1a and B1b cells,

while B2 cells include 2 subcategories of B cells -- marginal zone B cells and follicular B cells. These types of B cells serve different functions and can be found in varying levels in particular organs, and we attempted to profile these subsets across multiple organs. (APP).
In AD, the role of the adaptive immune system, and by extension B cells, is poorly. Moreover, considering the fact that AD is a neurodegenerative disorder, one particular tissue of interest is the meninges. Studies have shown that the lymphatic function of the 3 meninges influences AD pathology, thus making the meninges a particular tissue of interest in terms of its B cell activity and its relationship to AD [3].

We used two main techniques to analyze the B cell profile of AD mouse models. The first one is known as Fluorescence Activated Cell Sorting (FACS), wherein a machine takes a solution and establishes a flow where cells can be analyzed individually for particular wavelengths of light emitted by the cell. In preparation, cells are incubated with particular antibodies reacting to extracellular markers, and the FACS instrument detects the fluorescence and characterizes the cells in terms of the markers present. Using a particular panel of antibodies, B cells in the spleen, the meninges, and peripheral blood were quantified and subsetted.

Another technique used is known as immunofluorescence. In this technique, particular tissues are fixed with paraformaldehyde, and particular cells of that tissue are incubated with antibodies that bind to those cells. The antibodies are then attached with secondary antibodies bound to a fluorescent molecule, and the fluorescence can be examined under a microscope. The final image can be used to determine whether or not cells exist in a particular tissue, and if they do exist, their localization pattern can be derived from the staining. In our lab, the immunofluorescence was focused on the meninges, in an attempt to visually profile the local immune cell population.

Using FACS, we found that there were indeed significantly different levels of B cells between AD mouse models. In the future, we plan to use advanced statistical tools to determine the critical correlations present in the data that reveal which B cell subset is primarily implicated in AD pathogenesis. In terms of the immunofluorescence results, our data did reveal a varied population of immune cells in the meninges, the exact character of which remains to be seen in further experiments.

References

1. https://www.ncbi.nlm.nih.gov/books/NBK499922/

2. Kim K. *et al* (2021). Therapeutic B-cell depletion reverses progression of Alzheimer's Disease. *Nature Communications*, **12**(1), 2185.

3. Da Mesquita, S., et al. (2021). Meningeal lymphatics affect microglia responses and anti-Aβ immunotherapy. *Nature*, **593**(7858), 255–260.

MATHEMATICS AND COMPUTER SCIENCE

(L-R) Jacob Khalili, Benjy Katz, Ephraim Meiri, Yedidya Moise, Yehuda Snow, Yonah Moise

Finding Topical Similarity in Responsa Using Transformers

Jacob Khalili and Yehuda Snow

Advised under Prof. Jonathan Schler

One of the cornerstones of modern halachic literature is responsa. Careful learning, analysis and understandings of these scholarly letters inform not only the recipient on how they are to act, but establish important halachic precedence and principles used to decide complex questions for future generations.

Imagine a world where, when you are learning, helpful suggestions of sources to look at can help aid in your understanding; a world where it's simple and easy to find the opinions of other great poseqim (adjudicators of Jewish law) on similar issues and other cases where similar concepts are at play.

Using modern natural language processing (NLP) techniques, building such a tool on a large scale has only recently become possible. In 2018, Google Research published a state-of-the-art pre-trained network for NLP, called

Bidirectional Encoder Representations from Transformers, or BERT. BERT is an unsupervised machine learning model, which is used to create an encoding-vector which represents the texts. BERT's bidirectionality is one of the key reasons for its success. When representing a word, BERT not only looks at the word itself, but the context of the word, in both directions (previous and future words).

In April of 2021, researchers at Bar Ilan University published AlephBERT, a model based on Google's BERT, but trained on modern Hebrew literature. The model was trained on Twitter, Wikipedia, and OSCAR (a large dataset of multilingual data crawled from the internet.) The model was trained using Masked-Language Modeling (MLM), where every sentence fed to the model contained a blank word, which the model tried to predict; the weights of the model were optimized to predict the missing words correctly. This gave the model a sense of linguistic patterns of the language it was trained on.

Whether doing training or inference, the model requires a tokenizer to convert the words into a vector representation of the text. This vector representation is then fed into BERT to generate an encoding. When training, that encoding is then fed to a second half of the network, which predicts the missing word, followed by a loss calculation, and the updating of weights. When doing inference, the distance between the current paragraph's vector representations, and all other vector representations is calculated, and the closest paragraphs are represented to the user. After being trained in this fashion, the middle layer's output is used as a vector encoding of the input text.

In April of 2021, the Bar Ilan NLP Lab released AlephBERT, a BERT model designed for the Hebrew Language. It was trained on OSCAR's Hebrew section (a collection of articles in Hebrew from the internet), Hebrew Wikipedia, and Hebrew twitter. Because of the training data provided, the model learned very modern colloquial Hebrew. However, for use in this task AlephBert would need to learn a more formal, and much older Hebrew used in rabbinic responsa.

This was done by fine-tuned AlephBERT via MLM on a subset of the Bar-Ilan Responsa Project's collection of rabbinic responsa. Through this process, AlephBERT became accustomed to the dialect of Hebrew used specifically in responsa.

The model was then able to generate embedding in vector space that accurately represented the topics of each document. Calculating the magnitude of the distance between two vectors to represent how topically different they are. For computational efficiency, this process was implemented using matrix operations, which were performed on a GPU. Now an efficient method for comparing a paragraph to all the other paragraphs in corpuses of responsa and finding the paragraph that is the most topically similar to that paragraph was possible.

The results received were reasonable with this method, though the measurements are not objective. For any given paragraph about 2 of the top 3 suggestions were deemed relevant.

Generally in rabbinic responsa, the main points in the document are expressed in the beginning when the question is asked and at the end in the conclusion. So, in order to improve the search results document embeddings were also

generated. This was done by using the first 250 tokens and the last 250 tokens and in the document. This size was chosen because the maximum pre-trained AlephBERT model uses is 512 tokens.

This method on its own generated reasonable results which were similar in quality to the previous model. After summing both scores for any given paragraph (the sum of the paragraph distance and the document distance), when then sorted by overall distance, the results were promising, though similar. There is much difficulty in assessing model improvement unless there are very large changes, noticeable in reviewing a few dozen documents. In the future, once users are available, user feedback can be a critical tool in assessing how relevant a particular article is.

Overall, while the results for topical similarity were promising, there is still more work to be done. Sometimes results contained no paragraphs that were similar or just one or two similar results, even though there were actually many paragraphs that were similar. This could potentially be improved by increasing the number of training epochs, using an ensemble of models, and/or implementing a discriminator network.

Nikud Remover: Data Processing and Cross Data Type Compatibility

Benjy Katz

Advised under Prof. Jonathan Schler

The Bar-Ilan Responsa project was started in the 1960's as one of the first digitized, searchable databases in existence. The project is a collection of classical Jewish sources that are searchable and linked. This was revolutionary at the time for Jewish learning.

The problem in which I was tasked to solve is the resolution of the two conflicting Hebrew language standards: voweled: *nikud* (*chaser*) and non-voweled: non-*nikud* (*maleh*). In today's day and age, people exclusively write and type in *maleh*. Literature that is continuously added to the project is often written in *chaser*. The issue arises when a modern user initiates a query in *maleh* and requires that texts that are *chaser* show up as well.

Chaser Hebrew is written with small vowel points above, below, and within consonants i.e., וְצָנָנוּ. *Maleh* Hebrew is written without the small vowel points and instead substitute in other consonants i.e., וציוונו. The issue with converting from *chaser* to *maleh* is that there are no strict rules that the conversion adheres to.

The first way of addressing this problem is through a basic look up approach. Simply take every possible conversion from *chaser* to *maleh* and see which one is the most popular. This is done by keeping a large dictionary of words associated with their frequency in Jewish texts. For example, וְצָנָנוּ has four possible conversions to *maleh*: וציוונו, וצוונו, וציונו, וצונו The look up method simply checks which amongst these variations of the word is the most common. It decides to convert it to וצונו since it appears much more often than any of the other variations throughout various Jewish texts.

This method is a fair approximation at a valid conversion, but it does not consider many relevant factors. Simply by virtue of the fact that a word is common, does not necessarily mean that it is the correct conversion. A much more accurate method of conversion would be to construct a dictionary of *chaser* to *maleh* and

simply replace every *chaser* word with its corresponding *maleh*. In order to successfully implement this strategy, one needs a large dictionary with every *chaser* word and its corresponding *maleh*.

To build that dictionary, I relied on large corpuses of the same text written as both *maleh* and *chaser*. I was able to use this method as a type of Rosetta Stone to match corresponding *chaser* to *maleh* and enter that pairing as a dictionary entry. If the two corresponding texts matched perfectly to each other the algorithm to turn the two texts into a dictionary would be simple. Take the n^{th} word from the *chaser* text and the n^{th} word from the *maleh* text and that is a dictionary pairing, repeat for all n words and that completes the dictionary. Unfortunately, the texts that exists with two versions, a *maleh* and *chaser*, often have slight variations in how they cite sources and title paragraphs. As a result, the n^{th} word in the *chaser* might actually correspond with the $(n+5)^{th}$ word in *maleh*. If one would implement the simple matching algorithm with these non-perfectly aligned texts almost the entire dictionary would be incoherent.

To solve this problem, the two texts need to be simplified down to just what they have in common. Once the two text can be matched up exactly, aside from the fact that one is *maleh* and one is *chaser*, the simple matching algorithm can be used. The task of finding what it is that they have in common is also known as finding the longest common subsequence. This ensures that the algorithm preserves the maximum amount of good data that is available to it, while removing any small deviations between the two texts.

Using a naive brute force technique, the problem is np-hard and has a time complexity of $O(2^n)$ which essentially makes it non solvable with any reasonable sized data set and a modern computer. Using a dynamic programing approach, the problem is solvable in $O(n^2)$ time. The dynamic programing approach makes this problem solvable with large datasets, on modern computers, in a reasonable amount of time.

Now that a document with the deviations removed exists, the simple matching algorithm can be applied to match every *chaser* to a corresponding *maleh*. Once this dictionary has been created, converting a document from *chaser* to *maleh* is as simple as looking up every *chaser* word in the dictionary and replacing it with its *maleh* counterpart.

Paragraph Recognition from OCR Data in 18th Century Hebrew Books

Ephraim Meiri

Advised under: Dr Joshua Guedalia

Current OCR (Optical Character Recognition) methods allow for quick and cost-effective digitization of Hebrew texts, as the user must only check a small fraction of low confidence characters. Recently, as part of their larger effort to make rabbinic texts more accessible, Dicta – The Israel Center for Text Analysis – has developed a tool for digitizing books set in 'Rashi' style typefaces using such a method. While the OCR model does perform layout analysis in order to determine the flow of text, including complex multi-column layouts, the program's recognition of new paragraphs is lacking. Poor paragraph recognition is

problematic both for the end user reading the text, as well on the technical end. Being able to break the text into logical chunks is important for basic functions like storing, loading, or searching data.

Automatically finding new paragraphs is composed of two seemingly simple steps: 1) Identify visual/typographical features used in our set of books, and 2) search for those features in the text using the data produced by the OCR program. To briefly explain step #2: OCR converts an image to text by examining the black pixels, isolating a letter, and determining which letter is represented by those given pixels. In addition to the text generated, there is also location data stored in the form of a bounding box, indicating where each line is located. Since much of the information used to convey the beginning of a new paragraph comes in the form of whitespace, these bounding boxes become very important. In addition, some of the features used to indicate a new paragraph are simply not recognized by the OCR program. Thus, while paragraphs almost always begin with a visually distinctive headword, trying to identify these with bold detection or inference from size information was ultimately unsuccessful.

We Identified three features that could be used to identify paragraphs: 1) The second line of a new paragraph is indented, 2) new paragraphs often follow colons which are used to mark the end of paragraphs, 3) the last line of a paragraphs is often centered. While none of these metrics are sufficient on their own, together they allow us to identify new paragraphs with a high degree of accuracy. Accurately measuring these features is somewhat tricky. For example, while the size of the indentation is based on length of the headword above it, we found that comparing the size of the right margin to the rest of the page using a fixed ratio was actually more accurate than directly comparing the headword to the right margin. Unexpected results of this sort required careful examination of the data across an array of different books to make sure our method was both accurate and robust. Even while improving on preexisting methods, they remain useful as an estimate that then allows us to look at outliers in both the set of new-paragraph lines and regular lines.

After improving the overall results, we shifted to minimizing false positives. Narrowing our goal allows us to avoid checking whole pages— we could simply cut out the area around what was identified as a new paragraph. For this purpose, we automatically generated a html file that contained a random sample of lines identified as new paragraphs, with both an image cutout as well as debugging data for quick examination. We examined a seven-page sample, usually containing between 40 and 60 paragraphs, for around 170 books. Examining this sample of upwards of 7,000 examples allowed us to identify a number of recurring issues, stemming either from unusual page layouts or from errors in the OCR program's identification of line boxes or of the pages flow. After tweaking our code to avoid recognizing these errors, we had to examine a large sample of examples from our data to ensure that there were no (or a significant number of) correct paragraphs that look like (and be thrown out with) the errors we identified.

It is likely possible to fine tune our paragraph finder by examining even more samples, but we have likely reached the point of diminishing returns here. One front that might still benefit from further work is improved layout analysis for books with a mix of single and double

column text or with notes (*hagahot*) that split a column into two for a few lines. What we have learned about the book layouts can also be used to solve other problems, such as identifying headings.

Numerical Explorations in the Non-Linear Schrodinger Equation with Non-Symmetrical Gaussian Initial Conditions

Yonah Moise, Yedidya Moise

Advised under Professor Jeremy Schiff

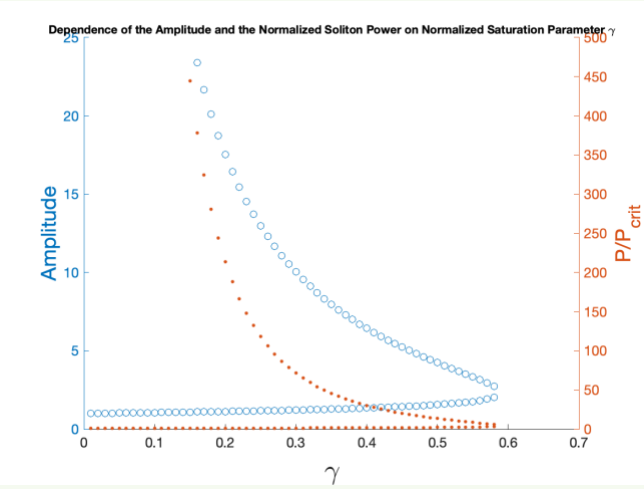
The Nonlinear Schrodinger (NLS) equation

$$-i\frac{\partial \psi}{\partial t}+\frac{1}{2}\left(\frac{\partial^2 \psi}{\partial x^2}+\frac{\partial^2 \psi}{\partial y^2}\right)+f(|\psi|^2)\psi=0$$

is a nonlinear partial differential equation (PDE) with physical applications in optics, fluid flow, and other fields. The nonlinear component of the equation can appear in many forms. When $f(|\psi|^2)=|\psi|^2$ the beam will self diffract if the power falls below the critical power, and will self-focus until collapse if above the critical power. This case is termed "Kerr non-linearity".

Instead, we focused on saturated non-linearity, with $f(|\psi|^2)=\frac{|\psi|^2}{1+\gamma|\psi|^2}$. In particular, we focused on the case of two transverse dimensions (x,y; spatial dimensions) against the dimension of propagation (expressed in the equation as t; the time dimension, this is the third spatial dimension z). Saturated non-linearity acts as a limitation on the nonlinear component of the equation to prevent it from blowing up as it $|\psi|$ grows.

We began our research by following the method laid out by Gatz and Herrmann [1] to solve for the spatial solution to the NLS. We applied the iterative method described in the paper to solve for the initial conditions (IC) of this PDE given our chosen parameters (γ) and constraints. We replicated Figures 1-3 from the paper to ensure our replication was sound. Our methodology was to run through a range of possible ρ_0 for each γ and to observe and pick the values which fulfill the constraint $\rho_1=\frac{1}{2}\rho_0$.



Most γ values had two corresponding ρ_0 which fulfilled the constraint. We used these ρ_0 values to find ρ vectors, which serve as the initial condition for this equation. Such an IC physically corresponds to the the initial state of the beam, which we use as a basis to step forward in time and approximate the development of the beam.

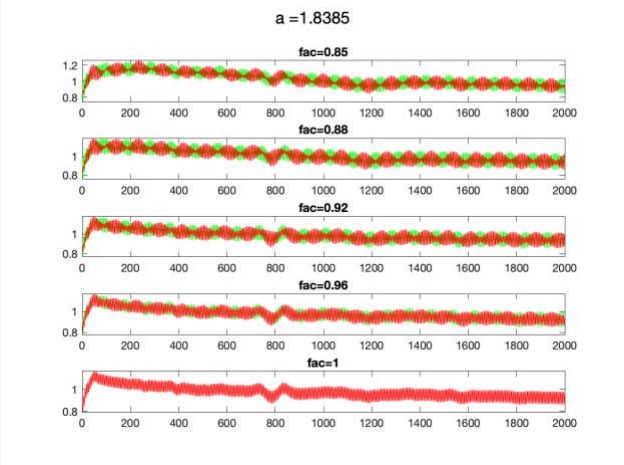
This leads to the second part of our research. We plugged our IC into a split-step method to step forward in time and approximate the stationary solution $\psi=e^{i\beta t}f(x,y)$ to our PDE for any given time.

We obtained the power of the beam from our soliton and constructed a Gaussian function of

two dimensions whose power is equivalent to the power of the solution. We ran the split-step method on the Gaussian approximation, plugging in different values for variables which affect the symmetry of the beam as well as the relative widths of the Gaussians. In all, we obtained data for 96 unique beams.

While running these methods, we collected data about how the amplitude and widths, in both dimensions, changed in any one beam over time. We observed that there are two internal modes relating to the widths of this beam.

Below is a figure showing the oscillations of the widths in both dimensions as the asymmetry decreases (when fac=1 the beam is symmetrical). We see a clear beating effect where the dimensions alternate dominant oscillations in their width.



In order to further analyze this data, we automated a process to take Fourier transforms of the amplitude and widths, and collect information concerning their frequencies. We summarize and present these data pertaining to the frequencies in tables.

Further research would involve running the codes with finer steps — we used 512 points on our intervals, but are interested in running the codes for 1024 and 2048 points. Further analysis must be done on the patterns that occur in the amplitude and widths as they change in the beam over time.

References

1. S. Gatz and J Herrmann, Journal of Optical Society of America B, **14**, 1795–1806 (1997).

ENGINEERING

(L-R) Chana Schwartz, Shana Erblich, Judith Wechter, Avi Radinsky

Measurements of contaminated water with low concentrations based on self-calibration point

Shana Erblich

Advised under Prof. Dror Fixler and Master's Student Alon Tzroya

Clean water is essential to sustaining life; contaminated water presents a very real risk to it. Contaminants in water can cause disease and death. Because of this, sensitive real-time detection techniques are necessary to ensure safe and clean water supply. Currently, water monitoring techniques are exclusive to the element they are detecting: each method must be tailored to the desired output. Most of these methods do not rely on optical properties, and each range of contamination in the classical methods of detection demand different

measurement techniques. We therefore suggest a new technique to measure water contamination using the full scattering profile (FSP) of a medium, wherein the intensity of scattered light at all angles around a medium is considered. We extracted an iso-pathlength (IPL) point that minimizes the effects of scattering and behaves as a self-calibration point. The IPL point is a crossing point between samples of different scattering coefficients (μ_s), and its intensity remains the same for a range of μs with a constant absorption coefficient (μ_a). The absorption coefficient does not change the angle of the IPL point, only attenuating intensity readings. Thus, it is possible to use the IPL point to directly measure the concentration of contamination in water in units of parts per million (ppm). In this paper, we show the appearance of the iso-pathlength point in single scattering regimes in concentrations of ppm. After building an optical setup to measure the angular intensity profile of variably contaminated water, we found a unique point for each sample diameter wherein light intensity remained constant irrespective of concentration. Additionally, we found that the IPL point varies linearly with sample diameter, and that absorption coefficients have no impact on the angular position of the IPL point. These findings suggest that the IPL point is an intrinsic parameter of a system which can be utilized in a variety of ways and mainly as an absolute calibration point. This method presents a novel means of efficiently measuring contamination in water.

Quantization of Neural Networks

Avi Radinsky

Advised under Prof. Adam Teman and PhD student Roman Golman

Over the past decade the Deep Learning revolution has taken the Computer Science research world by storm. Deep Learning is used in applications ranging from self-driving cars to protein folding to chess engines. But what is Deep Learning? How does it work? The full answer to the second question is currently unknown to the Computer Science community but perhaps if we answer the first question, we might gain a feeling for the answer to the second question.

The underlying algorithm behind Deep Learning is something called a neural network. For our purposes we can understand a neural network as a nonlinear algebraic equation. These are not like the equations that you dealt with in High School, rather they are massive equations that take thousands of inputs and have, at times, millions of coefficients. That isn't to say that these equations are complex, just very very big. So fine, we have a massive equation, how does that help me solve protein folding? This is the basic idea. You start off with an equation that produces absolute garbage for results, something that is totally random. Then you check to see what minor adjustments can be made to make the equation ever so slightly accurate. This check is done with a small batch of training data that you collect before. You then adjust the equation and check to see what adjustments can be made to make the equation just slightly more accurate. Rinse, Dry, and Repeat.

What comes of this is an equation that can be incredibly good at solving problems that it was trained to solve.

So where does my research come in? Remember how the equation that was created could have millions of coefficients? My question was to see if the number of unique coefficients could be compressed. For example: on a standard Deep-Learning equation, the coefficients can be a one of 2^{32} potential numbers. But maybe the equation doesn't require such a large range of coefficients to remain effective. Maybe one in 2^8 is effective, or one in 2^4, or even maybe one in 3. Even if this was effective, what benefit is gained from using a smaller range of numbers? The answer to that is simple, memory usage and speed. The smaller the number the less memory is used and the less time required to make computations. This is particularly important in a world where there are equations with as many as 175 billion parameters (GPT 3). If you could shrink the size of equations by even half, many hard to solve problems would now become within reach.

My research was to see what *shrinking* techniques were effective for retaining the accuracy of a well-trained equation. A number of approaches were designed and tested which involved shrinking the coefficient range to 2 and 3 and moderate success was found but it seems that shrinking the coefficient range while retaining a similar level of accuracy is a very hard problem that requires multiple iterations to get right. Resulting equations were far better than average but ultimately were not on par with current state-of-the-art.

Socially Pertinent Robots in Gerontological Healthcare

Chana Schwartz

Advised under Prof. Sharon Gannot and Mr. Pinchas Tandeitnik

Audio signal processing is the field concerned with processing and manipulating audio signals to refine and analyze the audio input. For example, in this lab they have used audio signal processing to filter out static from an old recording, to remove a noisy siren from the background of a phone call, and to pick out a single voice from a room of people talking.

To refine the audio source, an array of multiple microphones is used. When one is facing the array, the first microphone will pick up on the audio best, then the second and third, etc. Using this, the level of audio coming from different directions can be measured and analyzed accordingly. One project the lab is working on with this technique is to develop a better hearing aid. Each side would have an array of microphones to collect the audio. They are working on developing an algorithm based on signal processing and machine learning to extract and enhance the audio. Later on, they will be working on using the data collected from the two sides together to help the user be more aware of their surroundings.

The project I am working under is SPRING, Socially Pertinent Robots in Gerontological Healthcare. This project is being worked on by multiple universities across Europe, including 8 partners in 6 countries. The goal of SPRING is to create a robot (named Ari) which can be used in a gerontological hospital setting. Because the goal for it is to work among the elderly, it needs to fully be able to adapt to the people around

it, rather than the people adapting to it. The robot will need to be able to move around in a complex unstructured environment among multiple people, being able to see, hear, navigate, and react accordingly. To do this, SPRING needs to develop algorithms using computer vision, audio processing, sensor-based control, and spoken dialog systems using deep learning methodology, and then apply it to their robot and test it in a gerontological hospital setting.

Professor Gannot's lab is responsible for the audio processing level of the robot. This means that we are responsible for solely processing and analyzing the audio. The other universities are working on natural language processing (NLP), image processing, and developing social capabilities for the robot. Inside of the robot there is a circular array of microphones, which will be analyzed similarly to how the row array was described above. By analyzing the varied input from each microphone, we are able to filter out extraneous noise, pick out an individual speaking, and enhance the audio quality.

In order to test the audio system we are developing, our lab also wanted a virtual assistant to integrate into the system, which is what I worked on this summer. Building a virtual assistant involves a speech recognition system, the chatbot itself, and text to speech analysis to play the chatbot's response. There are two approaches available towards a speech recognition system. First, one can first record audio and then perform speech to text analysis on it. The second option is to transcribe the text as it is being spoken, which allows for a more seamless integration. This involves passing the data into a buffer, performing voice activity analysis on it, and transcribing each word as it is spoken. I did this using the Google Cloud Speech to Text API in Python.

Next, I needed to build the chatbot itself. I did this using Google Cloud's Dialogflow, which is a system to help build chatbots. This involved learning how to use the platform, training the chatbot, and integrating it via API in python. Lastly, I needed to convert the text responses from the chatbot to speech, which I did using the gTTS library. For future development the lab will need to convert my python script which was written on a Linux operating system to ROS, Robotic Operating System.

Gene Interaction Analysis with Abstract Boolean Networks and Formal Verification

Judith Wechter

Advised under Prof. Hillel Kugler and PhD student Eitan Tannenbaum

Every organism contains thousands of genes. These genes control the way the organism looks and behaves. But no gene exists in a vacuum. Many different genes interact with each other to create what are known as gene networks. Some genes activate, or turn on, others, while some inhibit, or turn off, the genes it interacts with. Understanding these networks and how the genes within them interact is of great interest to biologists. To understand how a network changes over time and how mutations of the system will affect it, one must first understand the network itself. This knowledge allows biologists to reverse engineer gene networks when necessary and understand the specific design and purpose of each one. It is possible to intuit network dynamics by hand, but as the number of genes

and interactions between them increase, so does the network complexity. To this end, it is possible and in fact advantageous to use computational biology to create a model of a gene-interaction-network. Many scientists even prefer to do this computationally, because it allows for more detailed explanations of what is happening within the system.

One well established method of representing gene networks computationally is that of a Boolean network. In such a network, each gene is either on or off. The interactions between genes are represented as signed, directed edges between genes, positive for an activating relationship and negative for an inhibiting one. Many networks also contain some interactions that scientists believe to exist based on experimental observation but are not certain about. To handle these interactions, Abstract Boolean Networks (ABN) are employed. In such a network each interaction is flagged as either definite (when scientists are certain that it exists) or optional (when there remains some uncertainty). A network can be modeled over time using update functions in which a particular gene's state is determined by evaluating a Boolean expression of the state's incoming nodes. Scientists can also add constraints to the model based on experimental observation. If the network was observed to be in a particular state at a specific time step, this is added to the Abstract Boolean Network to create a constrained Abstract Boolean network (cABN).

The "normal" way to determine a gene network that satisfies all constraints is to build several possible models manually and check each one against the constraints. Often, it is difficult and time consuming to find any models that satisfy all constraints. This causes many scientists to stick with simpler models that have fewer constraints to make it easier to satisfy them. Another problem with manual model construction is that it introduces bias from the person creating the model. For these reasons, scientists are very interested in an automated way of finding not just a satisfiable network, but all possible satisfiable networks.

The brute-force algorithmic approach to finding a gene network involves trying every possibly network until one is found that satisfies the constraints. This might work for very small models, but as more genes and interactions are added to the network, this becomes extremely unwieldy. The sheer number of possible networks means that checking every single one against the constraints incredibly time consuming. Using formal model verification has been successful in reducing runtimes to more feasible standards. The RE:IN tool uses the Microsoft Z3 solver to figure out which subsets of the optional interactions produce solutions that satisfy all constraints.

Sometimes, though, RE:IN finds no satisfiable solution. Often, this is because two or more of the experimentally observed constraints contradict each other. This can happen when the constraints are entered incorrectly into the program, or when the constraints themselves are incorrect and based off faulty scientific data. In other cases, RE:IN finds no solution because there is another interaction in the gene network that the scientist hasn't considered. Without it, the network has no solutions. In this way, RE:IN enables more of the genetic system to be discovered. The problem is that when there are no solutions, the RE:IN solver returns the message "No solutions found." but does not give any more descriptive information.

I developed a program that determines the Inconsistency Core of a model that was found to have no solutions. The Inconsistency Core contains the experimental constraints from any model that are inherently contradictory based on the definite interactions (also known as a Minimal Unsatisfiable Core). Once a scientist has access to the Inconsistency Core, he/she knows which set of constraints are preventing RE:IN from finding a satisfiable solution. In this way, a scientist only needs to verify the accuracy of a few constraints, instead of the entirety of the model. Scientist can focus his/her attention to understand what went wrong with the model. This cuts down on the time needed to fix a model before running it through RE:IN again and finding solutions towards gaining a biological understanding of the system.

CHEMISTRY, PHYSICS, AND LIFE SCIENCES

(L-R) Yosef Weiss, Yonit Krebs, Eliana Lebowitz, Bina Rosenblatt, Rivka Bella Rabaev, Leia Rubinstein, Simi Mirocznik (missing)
Abigail Razi

(L-R) Rebecca Aduculesi & Avigayil Roffe

Mapping the Molecular Basis of Protein E2F8 Using NMR Spectroscopy

Rebecca Aduculesi

Advised under Professor Jordan Chill and Dr. Inbal Sher

Regulatory proteins control the occurrences of various cellular functions. When a protein is regulated, it means that sometimes it will perform a certain function, while other times it will not perform this function, depending on the organism's needs. The more an organism can regulate its various functions, the more adaptive it can be, which makes it more efficient and powerful.

One common method of regulation in proteins is known as phosphorylation. When a ribosome creates a protein in the cell, it builds a specific sequence of amino acids that will then undergo various levels of folding to form the complete protein in its proper form. Out of the 20 amino acids, only serine, threonine, and tyrosine have a hydroxyl group. Phosphorylation occurs when a phosphate group reacts with a hydroxyl group, which means that only these three amino acids have the potential to be phosphorylated. When these amino acids are phosphorylated, the behavior of the protein can change. However, phosphorylation is a reversible process; the enzyme kinase adds phosphate groups, while the enzyme phosphatase cleaves off phosphate groups. Because the phosphate group can be taken on and off, phosphorylation is a typical mechanism for regulation.

The specific focus of this project is the protein E2F8, which belongs to the E2F family of transcription factors that are involved in regulating the processes in the cell cycle that determine whether cells will continue to mitosis or apoptosis [1]. However, the mechanism by which E2F8 operates remains unclear. Therefore, the ultimate goal of this project is to lay down the molecular basis for what happens in the cell, specifically focusing on the mechanism of E2F8. E2F8 has 8 different sites where phosphorylation can occur, and it was found by previous researchers that these phosphorylations seem to go in a specific order. This cascade of phosphorylations is another layer of regulation because a certain amino acid can only be phosphorylated if a separate amino acid is phosphorylated first. It is the goal of this project to map out the cascade of phosphorylations and to uncover how these various phosphorylations affect each other.

To accomplish these goals, an amino acid, threonine, at position 44 was changed to the amino acid, alanine. Alanine does not have a hydroxyl group, so this site no longer had the potential to be phosphorylated. The purpose of this mutation was to see if this phosphorylation has an effect that propagates to other, seemingly unrelated, phosphorylation sites. By using NMR (nuclear magnetic resonance) spectroscopy to observe the effect of this mutation, we can see if there is a connection between the sites. When the sample is loaded into the NMR spectrometer, the positions of the hydrogens and nitrogens are displayed, and based on this we can see whether a site has been phosphorylated. We can see the differences in the NMR spectrum with and without the mutation, and based on chemical shifts we can see if the phosphorylations of other sites were affected by the mutation. The NMR spectrometer allows us to see the specific position of the phosphates, which gives us insights into the mechanism of the protein.

To conduct the experiment, a different lab created plasmids with the various mutations and an antidote to a specific antibiotic. These plasmids were inserted into bacteria, and the bacteria was grown on plates that contained the antibiotic. Therefore, the bacteria were only able to grow on the plates if the plasmid was successfully inserted into the bacteria. This antibiotic selection allowed us to be sure that we would only collect colonies that contained the desired DNA. The colonies were then grown, and IPTG was added to make the cells start producing the desired protein. An SDS gel was run to show that the desired protein was being expressed (Figure 1). A miniprep was also done to extract the DNA out of the bacterial colony, and this DNA was sent to be sequenced. This sequence showed that the mutation was present because the codon for threonine was changed to the codon for alanine (Figure 2).

Once the experiment proved successful on a small scale, purifications were carried out on a large scale. To do this, the cells were first lysed and loaded onto a nickel column to extract the desired protein from the bacteria. Then the His-Tags that were still attached to the protein were cleaved using TEV enzyme, and the sample was loaded onto the nickel column again to separate the protein from the His-Tags. HPLC was then used to separate the peptides and the proteins. After this process, the sample was ready to be analyzed using the NMR spectrometer.

The obtained NMR spectrum (Figure 3) shows that we succeeded in making T44A which is an important reagent in the process of discovering the role of different phosphorylations and their order. The next step is to see whether this change influences the other phosphorylations.

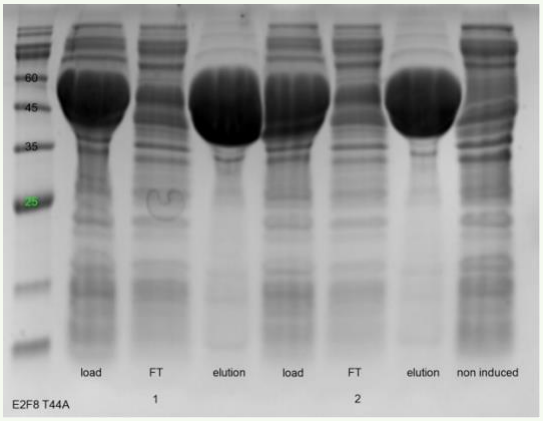


Figure 1. SDS gel showing expression of protein E2F8.

1. Wasserman, D. *et al*., Molecular Biology of the Cell, **31**(8), 725–740 (2020).

Nuclear Structure and Aging: The Relationship between lamin A/C, H3K9me3, and SIRT6

Yonit Krebs and Bina Rosenblatt

Advised under Prof. Haim Cohen and MSc student Ron Nagar

In the modern age, there has been a growing interest in the study of longevity—increasing human lifespan and reducing aging-related illnesses. Specifically, a growing body of research has indicated that sirtuin 6 (SIRT6), one of the seven mammalian sirtuins and an NAD+ dependent protein deacetylase that has several other functions, is associated with longer and healthier lifespans in mice. Overexpression of SIRT6 in mice has been shown to increase mouse lifespan and decrease some effects of aging such as genomic instability, inflammation, cancer and cognitive decline. In contrast, SIRT6-deficient mice begin

aging prematurely at two weeks and die within one month [1].

The deterioration of the cell nucleus is one place where one can see the effects of aging. Old nuclei lose the rigidity in the structure of their nuclear membrane, have greater DNA damage, and less heterochromatin. Lamin proteins are responsible for the nuclear membrane structure and interact with the lamin-associated domains (LAD) of the chromatin. Progeria, a rare disease of which one of the symptoms is premature aging, is the result of a mutation in the gene that codes lamin A/C [2].

Previous analysis of protein acetylation in the livers of SIRT6 mice found four lysine residues on the lamin A/C protein that had statistically significant higher acetylation rates: K97, K114, K171, and K180. However, SIRT6 is normally associated with deacetylation, presenting a puzzling question: Why do SIRT6 mice have increased lamin A/C acetylation? How does this effect the cell phenotype and what might be its connection to aging related disorders?

We used two different methods to investigate the relationship between nuclear structure and aging. First, we stained mouse liver tissue samples for tri-methylated lysine on histone three (H3K9me3) via immunofluorescence and visualized the results on a confocal microscope. Second, we planned to mutate the lamin A protein to be either acetylated or deacetylated at those four residues and transfect lamin A deficient cells with these mutated lamin A proteins and visualize the resulting phenotypes. As seen in Figures 1 and 2, both H3K9me3 and lamin A are associated with SIRT6 such that in SIRT6 knock-out (KO) cells, they are no longer expressed typically.

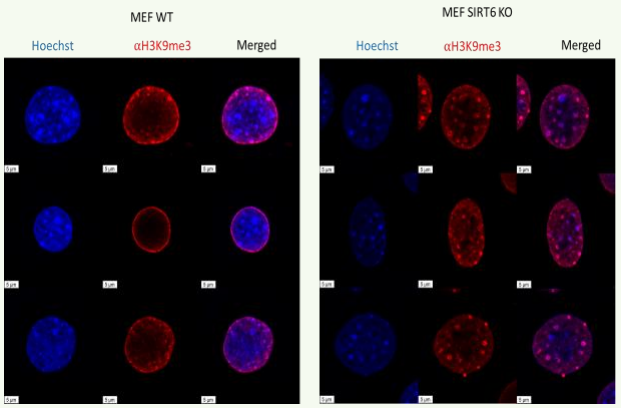


Figure 1. Relationship between H3K9me3 and SIRT6

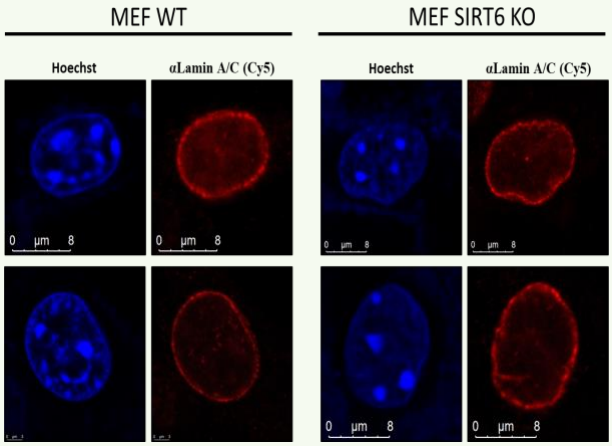


Figure 2. Relationship between lamin A/C and SIRT6.

H3K9me3

Previous research indicates that H3K9me3 is associated with the highly condensed sections of chromatin that make up the lamin-associated domains (LAD) [3]. This means that immunofluorescent staining for H3K9me3 can provide another way to visualize lamin, and the effects of aging on its interaction with the LADs. We preserved livers from young and old wild-type mice as well as young and old transgenic (SIRT6) mice in para-formaldehyde (PFA) and then used paraffin blocking and the microtome to cut the tissue samples into 8 µm slices that we placed onto positively charged slides. The

purpose of the slides' positive charge is this helps the negatively charged tissue sample adhere to the slide. We then dissolved the paraffin using xylene and rehydrated the tissue sample with ethanol and water. We calibrated an antigen retrieval method (TRIS buffer, 5 minutes) to break the covalent bonds between proteins formed during PFA fixation and retrieve the epitope. We then blocked the samples in Bovine Serum Albumin (BSA) with triton to prevent all unspecific protein binding and permeabilize the nuclear membrane. We incubated the samples with the primary and then the secondary antibody (conjugated to a fluorophore, Alexa) and stained the DNA with Hoescht as well. We used the Leica Stellaris confocal microscope to view the results.

Figure 3 shows the results of this experiment. Although, due to the limitations of our project, we were not able to fully analyze the data and quantify the results seen in these images, we were able to observe promising phenotypic differences among the four groups. First, the nuclei in the old groups were larger than those in the young groups, an expected result of aging. Second, while the H3K9me3 is localized to the nuclear membrane in the young WT and TG groups, the H3K9me3 in the old WT sample is more dispersed and looks somewhat similar to the SIRT6 KO phenotype observed in figure 3. However, the old TG group seems to have preserved some of the nuclear membrane localization of the H3K9me3, indicating the SIRT6 might play a role in this process. Ultimately, this technique can be used to observe the tissues of mice from other experiments and provide another way to look at their results. For example, we began this process for another experiment investigating the impact of exercise on SIRT6 expression and aging.

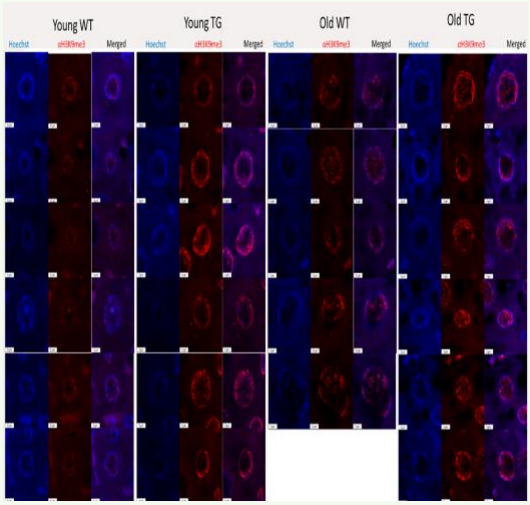


Figure 3. H3K9me3 immunofluorescence on liver tissue sample results.

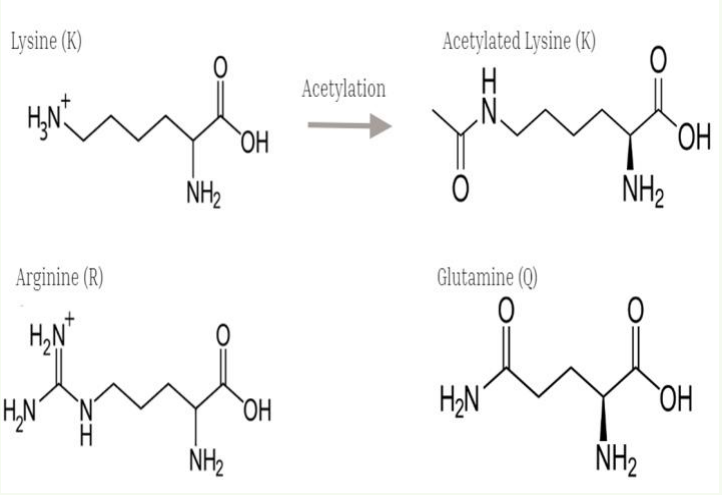


Figure 4. Lysine acetylation and replacement amino acids arginine and glutamine.

Lamin A/C Acetylation

As previously discussed, lamin A/C acetylation is associated with SIRT6 overexpression. To understand how lamin A/C acetylation might affect cell phenotype and aging, we mutated the lamin A/C protein in the four residues (K97,

K114, K171, and K180) to be either constitutively acetylated or permanently deacetylated. We mimicked permanent acetylation (see Figure 4) by replacing the lysine with glutamine (Q), an amide-containing amino acid. We imitated deacetylation by replacing lysine with arginine (R), the other positively charged amino acid. We confirmed a successful mutagenesis with Hy Labs sanger sequencing and then hoped to transfect lamin A/C KO cells with each of these plasmids to visualize how each residue's acetylation might impact the cell phenotype. Although, due to time constraints, we were not able to finish the transfection experiment, these results will shed light on the effects of lamin A/C acetylation on nuclear structure.

1. Kanfi, Y., Naiman, S., Amir, G. *et al.* The sirtuin SIRT6 regulates lifespan in male mice. *Nature* **483**, 218–221 (2012).

2. Karoutas, A., Szymanski, W., Rausch, T. *et al.* The NSL complex maintains nuclear architecture stability via lamin A/C acetylation. *Nat Cell Biol* **21**, 1248–1260 (2019).

3. Towbin, B., González-Aguilera, C., Sack, R. *et al.* Step-Wise Methylation of Histone H3K9 Positions Heterochromatin at the Nuclear Periphery. *Cell* **150**, 934–947 (2012).

Platinum-based catalysts for the oxygen reduction reaction in a hydrogen fuel cell

Eliana Lebowitz

Advised under Professor Lior Elbaz

The most common way of obtaining energy today is through the combustion of fossil fuels. Although fossil fuels are effective forms of energy, the byproducts of the combustion reaction include many pollutants that are harmful for the environment. In order to slow climate changes and preserve our planet, as well as our own health, it is important to find alternative ways of obtaining energy.

Current, "greener" alternatives, such as solar power or wind power, are intermittent, and we cannot rely solely on them to supply all of our energy demand.

In order to mitigate this issue, a new energy scheme has been developed: Hydrogen Economy. In this scheme, surplus energy produced during peak production hours from sustainable sources is used to produce hydrogen from water, which can be used later to produce clean energy. At the heart of the Hydrogen Economy is fuel cells technology. Fuel cells convert the chemical energy stored in hydrogen directly to clean electrical energy. They are very efficient and are eco-friendly as well– the only product of the reaction is pure water. However, it requires relatively large amounts of precious catalysts, mostly platinum-based, to carry out the hydrogen oxidation reaction and the oxygen reduction reaction. This hampers the wide commercialization of fuel cells. Through analysis of the properties of the platinum catalyst and the qualities that contribute to its efficiency, alternative, abundant and inexpensive catalysts can be tested to eventually be produced commercially to replace the expensive platinum, and realize the full potential of fuel cells.

In my project at Bar-Ilan University, I studied the properties of platinum-based catalysts for the oxygen reduction reaction, in order to better understand their behavior, and offer

alternatives. In this work, I synthesized Pt nanoparticles on a carbon support, using chloroplatinic acid and XC72. After the synthesis, the product was analyzed to determine the size of the particles and their surface area. A large surface area is an important attribute of a catalyst because the reactants which produce the electrical energy only come into contact with the surface of the catalyst. The larger the surface area, the more reactant that can come into contact with the catalyst and the less catalyst that is being wasted by being on the interior and therefore inactive. Figure 1 is the X-ray diffraction (XRD) spectrum of the platinum/carbon catalyst synthesized in this work. By comparing Figure 1 to the literature value for the XRD spectrum of platinum, the presence of platinum in the sample was confirmed, and using the Scherrer equation, the average size of the crystallites was determined to be 2.30 nanometers. Catalyst particles on the nanoscale have a large surface area, which indicates that it will be an effective catalyst.

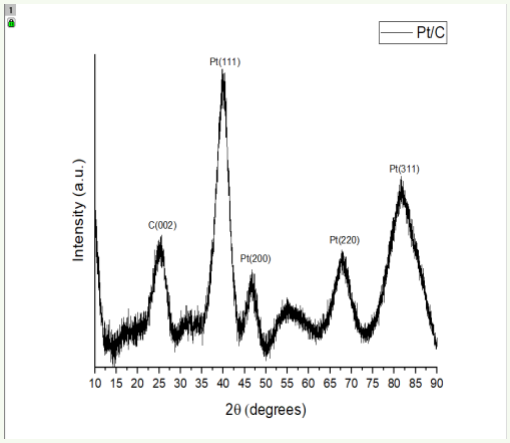


Figure 1. XRD spectrum of Pt/C.

After characterization of the sample through XRD, scanning electron microscopy and energy dispersive spectroscopy, the synthesized catalyst was studied in an electrochemical cell to determine its efficacy. Cyclic voltammetry was performed using the synthesized platinum over carbon catalyst to determine the electrochemically active surface area (ECSA) of the catalyst, which is calculated using the average of the area under the hydrogen adsorption peak and the area above the hydrogen desorption peak. The larger the ECSA, the more efficient the utilization of the catalyst.

Although cyclic voltammetry can produce a lot of information about the catalyst and the reaction, it is dependent upon the diffusion of the solution towards the electrode. The catalyst was also applied to a rotating disc electrode (RDE). Since the electrode is rotating in this system, the solution has a flux towards the electrode and therefore the diffusion becomes less and less of a limiting factor for the reaction to take place– this is illustrated in Figure 2. The higher the rotation speed, the higher the flux of the solution and therefore the higher the rate of the reaction and higher current is produced.

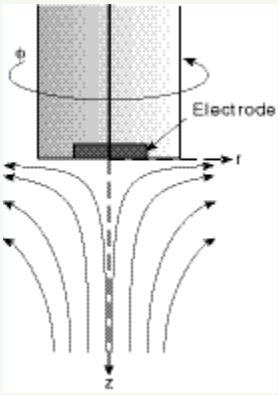


Figure 2. Illustration of flux of solution to rotating disc electrode

If oxygen is flowed into the RDE system, reduction will take place at the relevant potential range, and the effect of the rotation speed on the current can be observed. Figure 3 demonstrates that at 5 different rotation

speeds, the output current varies, increasing for each increase in rotation speed.

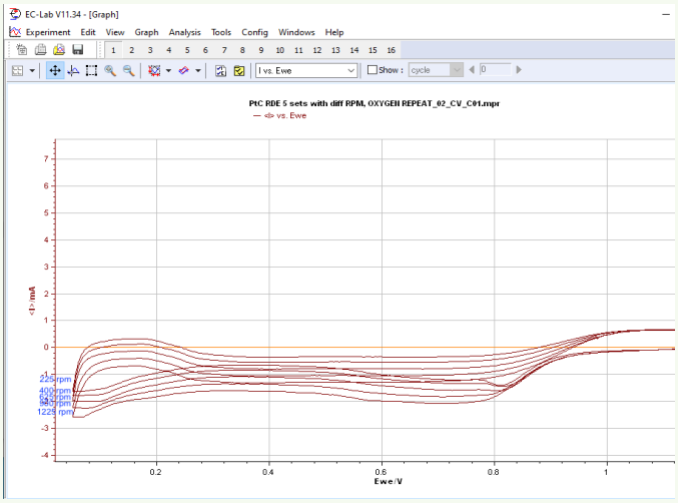


Figure 3: RDE with platinum/carbon catalyst and oxygen gas at 5 different rpm settings.

The number of electrons transferred in the reaction per oxygen molecule can be calculated using the Levich equation which incorporates the limiting current which is the maximum current measured at the given rate of diffusion, which is different for each rotation speed. For higher rotational speeds, there will be a higher limiting current.

After analyzing and testing a platinum catalyst and comparing the results to known results to ensure that the synthesis, analysis, and tests were carried out correctly, these methods can also be applied to other catalysts whose properties are unknown. These characterizations were carried out on multiple platinum-free catalysts as well.

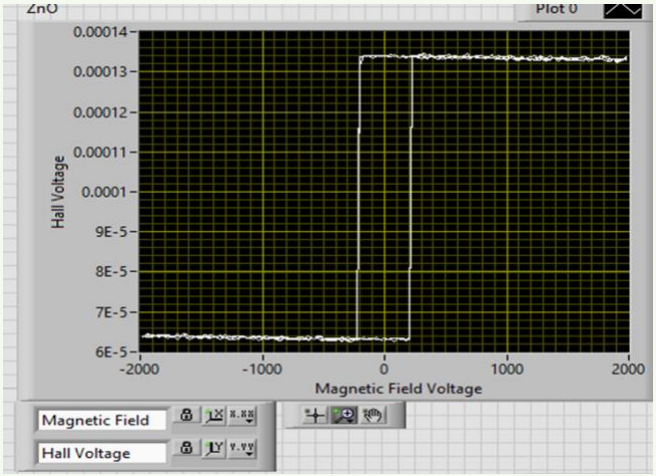
Electron Spin Interactions

Simi Mirocznik

Advised under Professor Amos Sharoni and Lidor Gery and Lior Saadia

The Hall Effect is an electromagnetic effect that occurs when a current passes through a material. This current induces a magnetic field, which exerts a force on the electrons, causing a buildup of negative charge on one side of the material. The difference in electric potential between the two sides is the Hall Voltage. A related phenomenon, the Anomalous Hall Effect, can occur in magnetic materials. When current passes through a magnetic material, it interacts with the spins of the electrons, magnetizing them to either be all spin up or all spin down. This spin interacts with the magnetic force and acts as a torque on the electrons, pushing them all towards one part of the sample depending on the spin direction and creating a potential difference – the Anomalous Hall Effect (ANE). In some magnetic materials, a magnetic field can cause magnetization switching where the electrons switch from being spin up to spin down, or vice-versa.

In order to observe the ANE and magnetization switching, solid state devices can be constructed via sputtering process. Ionized gas is accelerated towards the metal deposit knocking off atoms from the target metal to form a thin layer on a silicon substrate. Lithography is used to form electronic devices on the sample, which are then wire bonded to a PCB test board. This sample can then be tested in an electromagnet which applies a magnetic field, with current and voltage driven to it from Keithley source meters. The magnetic field can be perpendicular to the sample (out of plane) or parallel to the sample (in plane). Using the LabView software, the magnetization switching can be observed, and a hysteresis loop from the magnetization up and magnetization down will be formed.



I tested samples made of palladium and cobalt for magnetization switching. I also conducted measurements using a lock-in amplifier in place of the Keithley source meters, which generated an oscillating frequency of a specific amplitude in place of the applied current and voltage, respectively. I tested both the first and second harmonics of the frequency, and at varying amplitudes.

An Investigation of Post Transcriptional Splicing in Long Non-Coding RNA

Rivka Bella Rabaev

Advised by Prof. Yaron Shav-Tal and Chaya Bohrer

Introduction

RNA, Ribonucleic Acid, can be roughly categorized into two categories: coding and non-coding RNA. Coding RNA has long been the category primarily studied because of its role in the creation of all proteins essential for life. However, non-coding RNAs, specifically long non-coding RNAs, have been found to play essential roles in diverse biological processes such as proliferation, differentiation, apoptosis, migration and invasion [1–3]. Additionally, earlier studies have shown that the downregulation of lncRNAs are widely common in human tumors and a large number of lncRNA play an important role in cancer [4, 5], making them an important target for innovative cancer research.

Splicing is the process by which the introns are removed from the pre-mRNA and exons are joined together to create mature mRNA. In coding RNA, the identity of the protein can be altered depending on how many introns are removed, a phenomenon called alternative splicing. Alternative splicing is the cause of many diseases, including Hutchinson-Gilford Progeria Syndrome in which the lack of splicing in exon 11 of the LMNA gene causes the premature aging disorder. In non-coding RNA, alternative splicing can create various isoforms of a mature RNA which can interfere with the RNA's normal function. Splicing can occur at two points: co-transcriptionally (occurring while the RNA polymerase is transcribing the DNA) or post-transcriptionally (occurring after RNA polymerase transcribes the DNA and the RNA released from the site of transcription). Although the majority of splicing occurs co-transcriptionally, post-transcriptional splicing occurs in the transcript of interest (transcript x) in this study. Transcript x's expression levels are correlated with clinical stage and survival time for cancer x and has previously been found by the lab to undergo post transcriptional splicing in human cancer cells. This lncRNA can be used as a biomarker for cancer diagnosis and therefore might be a target for cancer therapies.

Research question

There were two research goals in this study. The first, to confirm that the lncRNA undergoes post-transcriptional splicing and the second, since the transcripts undergo post-transcriptional splicing, to understand how

splicing inhibition affects this transcript. Both the transcript and the cancer will remain unnamed and will be referred to as transcript x and cancer x throughout the abstract to protect the integrity of the study

Results

Confirming post-transcriptional splicing in transcript x. In order to detect the transcript in cells, we performed RNA FISH, a technique that allows the visualization of target RNA, using specific probes that hybridize with the RNA transcript. When we directed probes that hybridize with the exon and intron transcript, we observed unspliced transcripts (exon and intron overlap) not only at the transcription site, but also in the nucleoplasm. If co-transcriptional splicing had occurred, we would have expected to see an intron signal only at the transcription site, indicating that splicing was happening exclusively at the transcription site. Instead, unspliced transcripts in the nucleoplasm confirmed that post-transcriptional splicing was taking place.

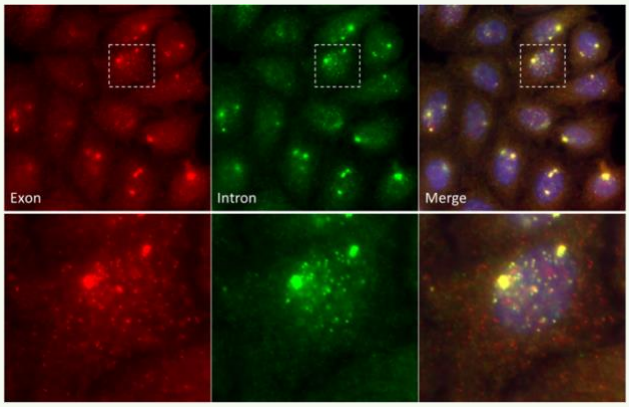


Figure 1. The lncRNA transcript undergoes post transcriptional. Unspliced transcripts were observed in the nucleoplasm. Small dots are RNA transcripts. Large dots are sites of transcription. RNA FISH exon (red) intorn (green). Hoechst DNA stain is in blue. Bar = 10 µm (bright dots in cytoplasm for exon, bright only in nucleus for intron- immature transcripts- but not only in transcription site)

To confirm these findings with quantitative evidence, we performed a PCR with specific primers directed to different regions of the lncRNA transcript. If co-transcriptional splicing had occurred, the PCR should have only produced bands in the wells for the exon, since the intron undergoes degradation. Instead, the PCR showed a strong band in the intron well, as well as two bands in the exon well. These findings quantitatively confirmed the presence of post-transcriptional splicing in transcript x as well as the presence of multiple isoforms.

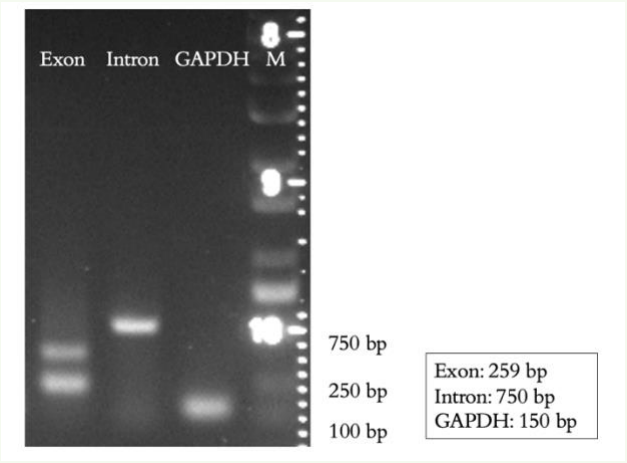


Figure 2. PCR confirms that the lncRNA transcript undergoes post-transcriptional splicing. The band in the intron well indicates the excitation of unspliced RNA, which quantitatively supports the thesis of post transcriptional splicing. Two bands in the exon well indicate presence of isoforms.

Examining the effects of splicing inhibition

We treated cells with Pladienolide b (PLB), a splicing inhibitor, and conducted an RNA FISH to examine what happens when the introns remain in mature RNA. Although we expected the treated cells to show more transcripts with introns (since no introns were removed) as

shown in Figure 3, the untreated cells have far more transcripts present. This indicates that the splicing of introns may be crucial to the overall survival of lncRNA.

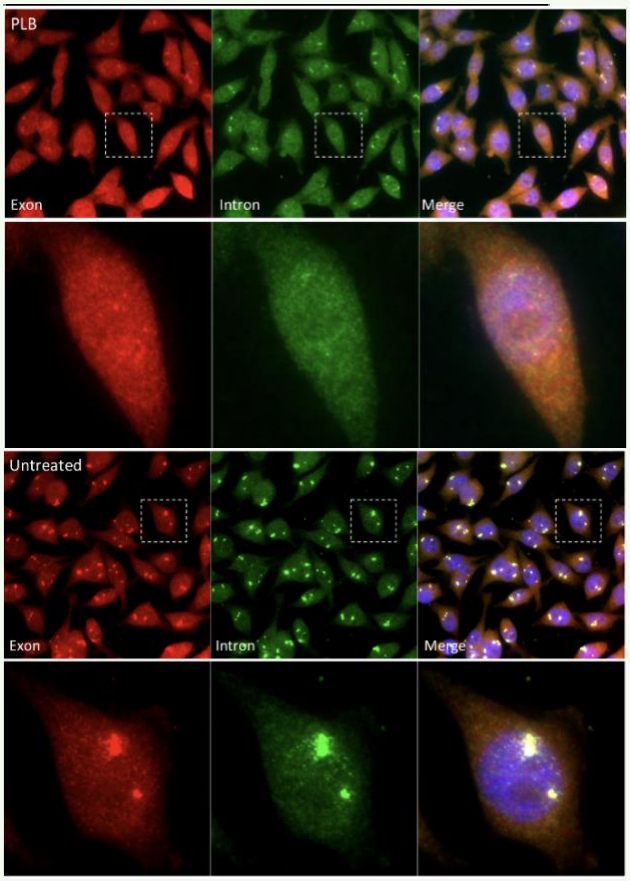


Figure 3. Splicing inhibition decreased lncRNA transcript expression (right). Untreated cells have higher expression of lncRNA transcript than the (Left) treated cells. RNA FISH exon (red) intorn (green). Hoechst DNA stain is in blue. Bar = 10.

Using software called Imaris, we then conducted a quantitative analysis on the number of cells with transcription sites compared with the number of cells without transcription sites. As seen in Figure 4, the graphed data clearly indicates that the number of cells with transcription sites severely decreased when cells were treated with a splicing inhibitor, indicating that splicing inhibition decreases the survival of lncRNA transcripts.

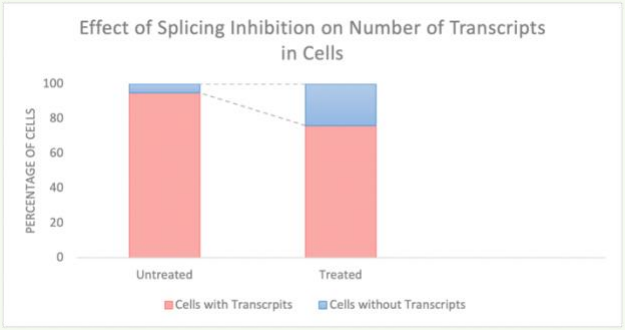


Figure 4. Splicing inhibition decreases lncRNA transcripts expression at TS. Quantitative comparison of cells with TS in treated and untreated groups.

To confirm these results, we performed a knockdown (siRNA) of SON, which is a core protein of the nuclear speckles. Nuclear speckles are sub-nuclear components that contain splicing factors. The results were consistent with the previous results where we added a splicing inhibitor. As seen in Figure 5, expression of lncRNA decreased after treatment.

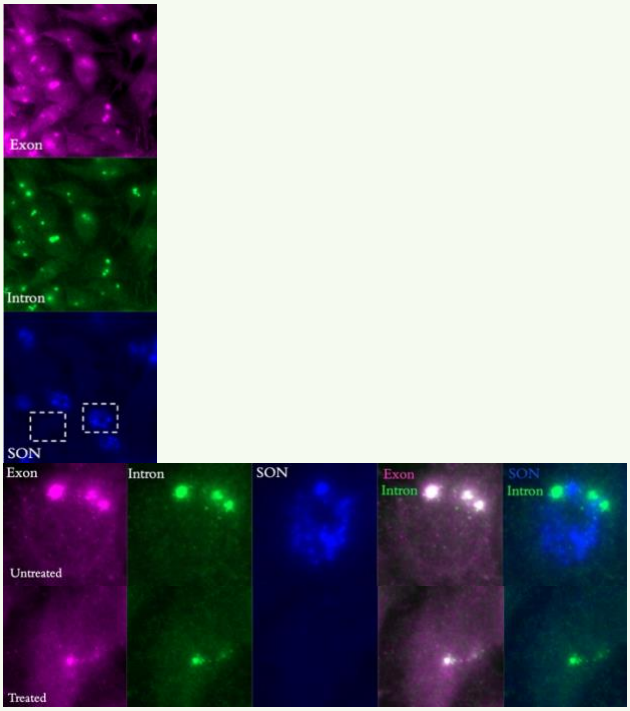


Figure 5. SON depletion results in lncRNA transcripts decrease. Knockdown of SON RNA FISH exon (magenta) intron (green) SON (blue). Bar = 10.

Conclusions

These findings indicated that transcript x undergoes post transcriptional splicing, multiple isoforms are present of transcript x, and under splicing inhibition there was a decreased expression of lncRNA transcript. Future directions include understanding if and how multiple isoforms are involved in the upregulation of cancer x as well as further research on the importance of introns (or the lack of introns) to the survival of the RNA.

1. Wu, X., Zheng, X., Cheng, J., Zhang, K., and Ma, C. (2020). LncRNA TUG1 regulates proliferation and apoptosis by regulating miR-148b / IGF2 axis in ox-LDL-stimulated VSMC and HUVEC. Life Sci. **243**, 117287.

2. Zhang, Z., Li, J., Guan, D., Liang, C., Zhuo, Z., Liu, J., Lu, A., and Zhang, G. (2018b). A newly identified lncRNA
MAR1 acts as a miR-487b sponge to promote skeletal muscle differentiation and regeneration. J. Cachexia, Sarcopenia Muscle **9**, 613–626.

3. Zhao, W., Geng, D., Li, S., Chen, Z., and Sun, M. (2018). LncRNA HOTAIR influences cell growth, migration, invasion, and apoptosis via the miR-20a-5p/HMGA2 axis in breast cancer. Cancer Med. **7**, 842–855.

4. Adams, B.D., Anastasiadou, E., Esteller, M., He, L., and Slack, F.J. (2015). The Inescapable Influence of Noncoding RNAs in Cancer. Cancer Res. **24**, 5206–5210.

5. Peng, Y., and Croce, C.M. (2016). The role of MicroRNAs in human cancer. Nature **28**, 15004.

Computational Modeling Cell Wall Protein

Avigayil Roffe

Advised under Professor Hanoch Senderowitz

Proteins are specific in their jobs and crucial for life. Diseases stem from a lack, build up, or error in a protein function. Accurate and precise modeling of proteins provides information necessary to understand what the protein does and how the protein does it. Building the cell wall of an organism requires important and specific proteins.

I worked on a project funded by the European Union to find small, environmentally safe, and effective pesticides. This will increase food security on a global scale. In order to achieve this goal, computational chemists help model and predict what a good pesticide would look like. To find a good pesticide, scientists look into the pest. I looked into how the bacteria of

a pest was forming the cell wall, understanding that a mutation in this process would kill the pest. I used different software to predict potential 3D structures of a protein. Once I had these models, I was able to simulate millions of potential ligands that could act as a potential pesticide. These pesticides would inhibit cell function and terminate the protein.

Pests can destroy entire fields of crops. I set out to model cellulose synthase catalytic subunit. The first step was to identify the binding sites and mechanisms of cellulose synthase in this enzyme. Next to choose the best site to inhibit, and then run a software to find potential ligands that would inhibit this protein.

The binding sites and the mechanism of the query protein were found by searching for homologous proteins. This is called homology modeling, and sheds light on a protein's binding patterns via a potential secondary structure. These proteins found have similar sequence identity and common ancestral history. I read the papers on the four homologous proteins of my query protein that described the binding pattern and mechanism.

From these papers, I discovered a likely mechanism the query protein use. The reactants of cellulose synthase are sugar, UDP, and cyclic-di-GMP. The product of the reaction is cellulose. For each reactant, there is a unique binding site on the protein. The protein has three important binding sites. For this project, I checked out the site that binds to UDP. This site was contained in the amino acid sequence I was modeling, and I was excited to investigate more and come up with a potential structure to the amino acid. There are two complexes in play: BcsA and BcsB. The finger helix contacts the polymers terminal glucose in the BcsA gating loop. Once the sugar is looped, the helix moves up and down – the substrate is binding, and the polymer elongates. This pushes the polymer into BcsA's transmembrane channel and connects the sugar to the cell wall resulting in cellulose.

Next, I modeled the protein. Protein modeling is a technique used to predict the structure of a protein. Protein modeling takes one or more template proteins and a target protein. The template proteins were the chains I identified on the homologous proteins that were similar to my query protein. The target protein was the query amino acid sequence.

This code ran and resulted in a pdb file of a protein. A pdb file lists every atom with its x, y, and z coordinates in the protein. There are some challenges with computational modeling. It is difficult to predict the secondary structure of a protein. Proteins are long chains of amino acids that are flexible with the ability to rotate; they can make infinitely many conformations. Additionally, there are the conditions the protein is in that may promote or discourage a certain conformation. The method isn't perfect, but will allow us to reduce the number of potential ligands.

I was able to form five potential models of the protein using a software called Modeller. Modeller takes in an alignment of the template and target sequences. After the alignment file is inputted, five models are made and run through different statistical tests. The results are five output files aligned with the new pdbs of the model. There is the DOPE assessment score, the Modeller objective function, or the GA341 assessment score.

This project continues by researching the other two binding sites and seeing if a mutation in

that binding would inhibit the cellulose synthase. The sugar is not the only thing the protein binds to. UDP, urine diphosphate, comes in as an activator. Once the models are made, they are tested against billions of potential ligands. This helps identify potential pesticides that need to be tested in a wet lab.

Using the Optogenetic Organ Gut Culture System to Study Neuro-Immune-Microbiome Crosstalk in the Gut

Leia Rubinstein

Advised under Dr. Nissan Yissachar and PhD student David Jessula Levy

The gastrointestinal tract is densely populated by microorganisms including, bacteria, fungi, viruses, protozoa and archaea, collectively known as the gut microbiome. Previous research has established the critical role the gut microbiome has in an individual's overall health and well-being including assisting in digestion, synthesizing nutrients and developing the immune system. In addition to the microbiome, the gut also contains many immune cells and neurons of the enteric nervous system, a subset of the autonomic nervous system that controls peristalsis. To maintain homeostasis, the enteric nervous system, microbiome and immune system need to remain in perfect balance. Changes in this equilibrium are associated with inflammatory bowel disease, cancer and pathologies of the central nervous system [1].

It has been already determined that neurons are affected by the microbiome, immune cells and epithelial cells in the gut [2]. However, the specific role of the neurons in the gastrointestinal tract is a topic that has been poorly explored. This research seeks to determine the function of neurons inside the crosstalk of the immune system and microbiome, specifically focusing on cholinergic neurons, one of the most common neurons which contains the enzyme choline acetyltransferase (ChAT).

Optogenetics refers to neuronal activation by a light sensitive ion channel. It is a method for triggering the response of specific neurons, as opposed to agonists and antagonists which are nonspecific. In this experiment, blue light was used to open the light-gated ion channels, channelrhodopsin, of the cholinergic neurons, causing an influx of ions and activating the mutant cell.

Mice with the ChAT/Cre genes were bred with mice containing the ChR2/YFP genes, generating offspring with ChAT/Cre-ChR2/YFP, where only cells that are ChAT positive will be activated by blue light. The Cre protein cleaves the lox-p region, taking out the stop codon, allowing the ChR2-YFP expression.

The experiment was carried out using a modified version of the gut organ culture system (Figure 1), a unique device created by Dr. Nissan Yissachar that facilitates the studying of an entire colon without causing damage and is easier to manipulate than an *in vivo* model. This device allows six intact colons to be placed in monitored conditions to study the gene expression after treatment with bacteria, antibiotics etc. and observe the interaction between the microbiome and immune system [3].

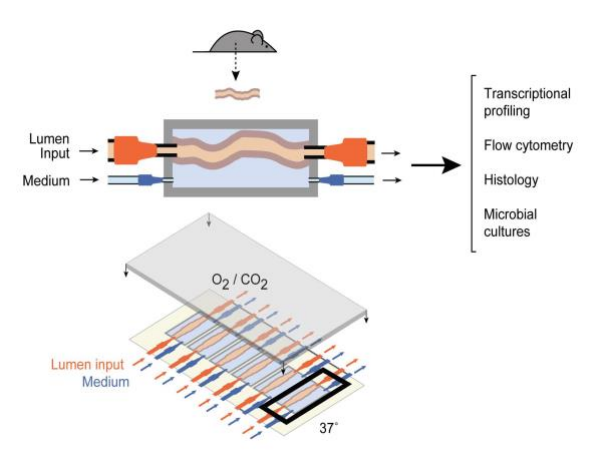


Figure 1. Gut organ culture system. Intact colons are connected to the input and output pumps of the device that control the flow of the medium into the gut lumen. Gas outlet in the lid allows the flow of an oxygen and carbon dioxide mixture into the device.

The gut organ culture system was improved to create the optogenetic organ gut culture system that has all the advantages of the original system with the addition of a software that has the ability to control the stimulation of neurons (Figure 2).

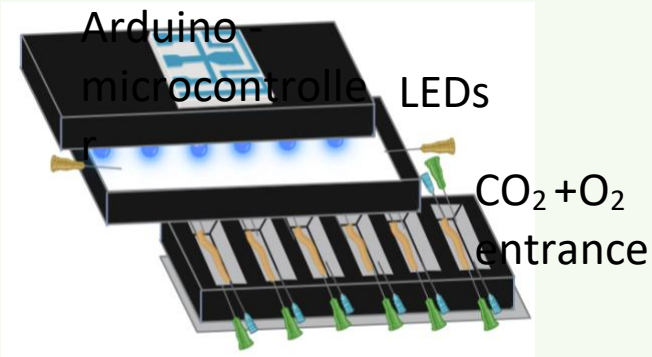


Figure 2. Optogenetic gut organ culture system. An improved version of the original gut organ culture system that incorporates a microcontroller with blue or yellow light embedded and a software that enables the control of light stimulus.

To validate the new system, the gut from mutant mice was connected to the device and triggered with blue and yellow light. An increase in fluorescence signal of cFos was detected after blue light stimulus compared with yellow light stimulus, indicating that blue light can activate the neurons in this system (Figure 3).

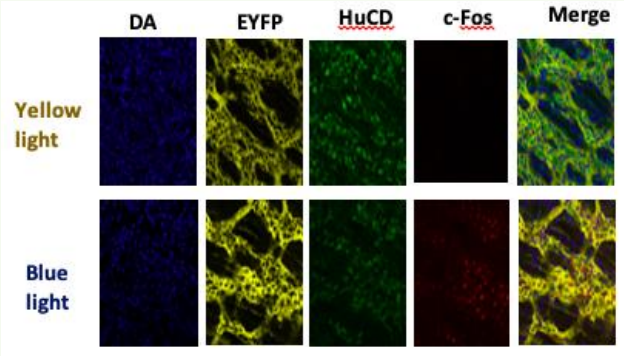


Figure 3. Imaging analysis showed an increase of cFos intensity in tissues that were triggered by blue light when compared with yellow light stimulus. Tissues were fixed with 4% PFA and stained with: DAPI (Blue), HuCD (green), EYFP-endogenous (yellow) and c-Fos (red). Tissue was stimulated through yellow and blue light.

RNA extraction from the whole tissue was performed to determine the consequences of neuronal activation from the blue light. Then, using Real Time PCR, it was possible to check the expression of selected genes after blue light stimulus as compared with yellow light stimulus. The results showed an increase in proinflammatory cytokines IL6, IL17, neuropeptide VIP and epithelial cells ICAM (Figure 4) indicating that activation of the neuron modulates many processes in the tissue.

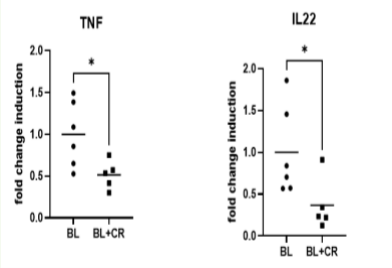


Figure 4. Transcriptional changes in the whole tissue after blue light activation through RT-PCR. Blue light stimulation showed greater activation of genes IL6. IL17, VIP and ICAM as compared with yellow light stimulation.

A preliminary experiment is now taking place to incorporate the bacteria *C. ramosum* (known to repress many genes) into the experiment to determine the effect neurons have on the gut microbiome. Whole mount samples of the gut both with and without the bacteria *C. ramosum* were introduced into the optogenetic organ gut culture system. The results showed that the proinflammatory cytokines such as TNF alpha and IL22 are suppressed in tissues co-cultured with *C. ramosum even* though they were stimulated with blue light (Figure 5). This research is still ongoing but the results are promising that neuron activation in combination with *C. ramosum* bacteria may play a role in limiting the immune response in the gut.

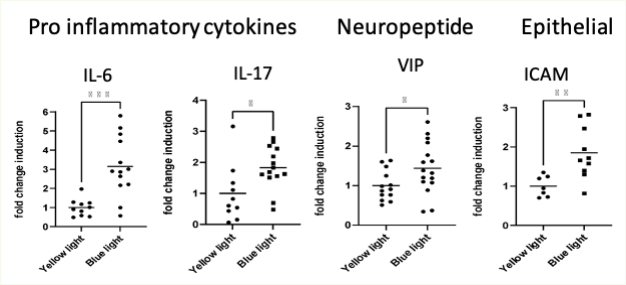


Figure 5. Transcriptional changes in the whole tissue co-cultured in *C. ramosum* after blue light activation through RT-PCR. Tissues co-cultured in *C. ramosum* showed suppressed activation of the proinflammatory cytokines TNF and IL22 following blue light stimulation.

While the crosstalk between the gut microbiome, enteric nervous system and immune system is a topic that has been previously studied, the specific role of the neurons in this crosstalk has yet to be extensively researched. Through the use of the optogenetic organ gut culture system, the effects of the neurons can be uncovered, paving the way for further research into the specifics of the communication in the gut.

1. Yissachar N, Zhou Y, Ung L, Lai NY, Mohan JF, Ehrlicher A, Weitz DA, Kasper DL, Chiu IM, Mathis D, Benoist C. Cell. **168(**6), 1135–1148.e12. (2017).

2. Yissachar N. Curr Opin Neurobiol. **62**, 26–33. (2020).

3. Azriel S, Bootz H, Shemesh A, Amidror S, Yissachar N. J Vis Exp. **172**, e62779 (2021).

Three-dimensional nanoparticle assembly by a modulated laser-induced microbubble for fabrication of a micrometric pattern

Yosef Weiss

Advised under Prof. Hagay Shpaisman and PhD student Udi Greenberg

During the summer of 2022, I was privileged to work in the Directed Materials Assembly laboratory in the Nanotechnology Department of Bar Ilan University. I joined a project

exploring the use of laser-induced microbubbles for both two-dimensional and three-dimensional printing. The underlying mechanism requires focusing a laser beam at the interface between a substrate and a liquid containing nanoparticles or ions. Due to laser heating, microbubbles form and convection currents pin deposits at the liquid/bubble/substrate interface [1].

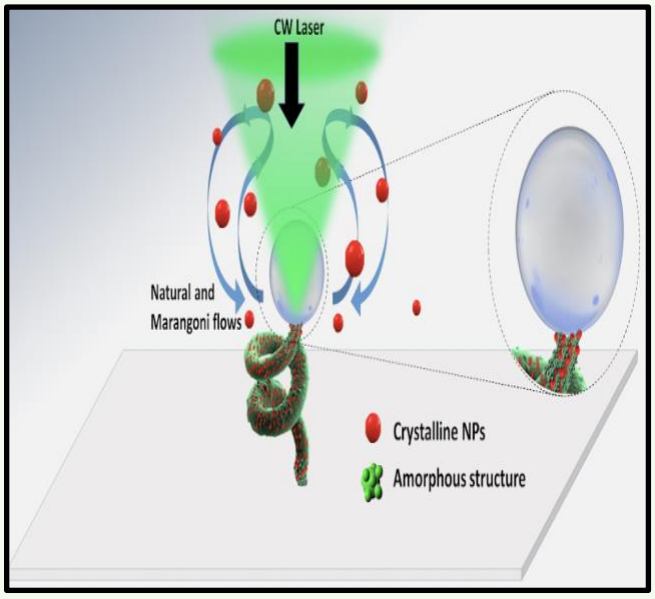


Figure 1. Laser-induced microbubble three-dimensional printing mechanism

Through the last nine years, this lab has developed a reliable method of two-dimensional, nanoparticle laser printing. Benefits of this printing format include achieving extremely narrow and precise lines (on the order of nanometers), the ability to print with a wide variety of nanoparticles, and a seamless transition between many different nanoparticle chemistries. With this technology, users can also stack multiple electrical connections by alternating between conductors and insulators.

When expanding this method to three-dimensional printing, two different techniques emerged—layer-by-layer printing and vector printing. Layer-by-layer printing builds up a design similar to a traditional three-dimensional printer; albeit using the laser technique. Vector printing directs the laser onto a continuously growing "stack" of material. For each of these printing styles, an issue arises. (1) For layer-by-layer printing, inconsistent deposition occurs due to sinusoidal patterns in the previous layer. The origin for this phenomenon is a self-enhanced process where areas with slightly more deposits tend to absorb more light, thus leading to even more deposits at the same location with each additional layer. (2) For vector printing, when the laser deviates from the vertical, pillar-like structure, deposits tend to accumulate unevenly at the bottom of the microbubble leading to deformation of the microstructure.

When approaching the former issue, the first step is to determine the optimal laser parameters for three-dimensional printing. Parametric properties include laser intensity, frequency, velocity, and duty cycle. Once two quality sets of parameters were identified, multiple layered lines were printed in order to visualize the pattern of deposition in the "wall" microstructure. Using a profilometer (optical probe), these layers were characterized in their height and roughness. It was previously determined that a shorter laser residence time results in less deposition due to a shorter period of nanoparticle heating [2]. To solve the problem of inconsistent deposition in layer-by-layer printing, a modulating laser was programmed to speed up and slow down depending on the height of the previous deposition. Through analyzing multiple

examples, the period of the pattern of deposition was determined and then used for the modulation of the laser to correct (smooth) the deposition pattern.

For the latter issue, a double-axis stage was used to allow the nanoparticle solvent to be rotated to any orientation in all three dimensions. By rotating the stage, it keeps the "pillar" always parallel to gravity and therefore does not allow the structure to become unstable and irregular. A casing for these glass slides was designed and printed in-house in order to prevent the glass slide from falling off during rotations. A program was then written in order to calculate the degree of rotation for both axes and control the rotating stage to create the more complex shapes.

Developing the understanding and innovative tools for creating 3D structures with nanoparticles has a number of potentially significant applications. Some examples include creating compact stacked computer chips not possible with current technology, precision micro-motors and other mechanical devices, and micro-probes for visualization and measurement. The unique combination of micro/nano-size and high-precision will likely also lead to unanticipated novel applications.

1. E. Greenberg *et al*., Adv. Mater. Interfaces, **6**, 1900541 (2019)

2. N. Armon *et al.*, ACS Appl. Mater. Interfaces, **9**, 44214-44221 (2017)

Congenital Dyserythropoietic Anemia Type I (CDA 1); Exploring the Properties and Interactions of Codanin-1

Abigail Razi

Advised under Professor Benny Motro and PhD student Moriya Kahta

Congenital dyserythropoietic anemia type I (CDA I) is an autosomal recessive disease associated with abnormalities in erythrocyte precursor proliferation that may be expressed as spongy heterochromatin. The disorder is mostly caused by a mutation in the CDANI gene which encodes for Codanin-1 protein. Codanin-1 is a conserved protein which serves as a scaffold for proteins and is presumably involved in chromatin assembly and maintenance.

Our research focuses on the study of the function of the Codanin-1 protein. Previously, it was demonstrated that Codanin-1 acts as a protein scaffold to recruit the histone chaperon protein Asf1. Overexpression of Codanin-1 results in a shift in the localization of Asf1 from the nuclear region to the cytoplasm.

In a subsequent analysis of the interactions between Codanin-1 and Asf1, the possible influence of Codanin-1 on post translational modification (PTM) of Asf1 was studied. To do so, HeLa cells underwent an inducible knockout of the gene that codes for the Codanin-1 protein. The prediction was that the knockout of this gene would prevent the expression of Codanin-1, thereby influencing the activity and PTM of Asf1 as well. Then, immunoprecipitation was conducted to precipitate Asf1 from treated and untreated cells. Mass spectroscopy was used to identify PTM in Asf1 from cells containing endogenous Codanin-1 and cells containing the inducible knockout of the Codanin-1 gene. Differences in PTM between the endogenous Asf1 protein and the Asf1 protein following the inducible

knockout of Codanin-1 would not only allude to the influence of Codanin-1 on the activity of Asf1, but more importantly, it may point to larger revelations regarding the role of Codanin-1 in cellular function. The results of this study regarding Asf1 modifications have yet to be identified by mass spectroscopy.

In an additional project, I aimed to change the endogenous Codanin-1 gene by introducing a construct that contained a degron, just prior to its stop codon. A degron is a portion of a protein that is important for controlling protein degradation at the proteasome. In this study, the degron is Auxin (IAA)-dependent, namely in the presence of IAA receptors in the cells. When IAA is added, it binds to its receptor. The receptor recognizes the degron and sends Codanin-1 to the proteasome to be degraded.

To insert the plasmid into the genome, a construct was created (referred to as #1217) by adding a 5' arm and a 3' arm in order to allow for homologous recombination to occur. Crossover then occurred between the endogenous chromosome and the construct containing the degron and Hygromycin antibiotic resistance. The resulting chromosome contained the same DNA as the endogenous Codanin-1 gene, but also contained the additional degron and antibiotic resistance as well.

Because there are two chromosomes, it is necessary to introduce the degron into both strands using two constructs that confers cells resistant to both antibiotics: Hygromycin and Blasticidin. The introduction of this gene into the chromosome induces antibiotic resistance by producing a protein that degrades its respective antibiotic, thereby deeming the cell resistant. Since cells will only survive if the antibiotic resistance has been introduced, cell survival would indicate which cells have been successfully transformed. In cells in which the degron was inserted to both alleles, the addition of Auxin should lead to the degradation of Codanin-1.

Using Crispr methodology, we utilized an sgRNA to cut a strand of chromosomal DNA in order to insert the degron which contained Hygromycin resistance. To insert the degron into the second chromosome, we looked to excise the Hygromycin antibiotic resistance fragment of the plasmid between NheI and NheI and replace it with Blasticidin antibiotic resistance. In order to insert the Blasticidin antibiotic resistant, we edited vector #1217, represented in Figure 1, by inserting a fragment from plasmid #1158 (Blasticidin resistance) in place of Hygromycin resistance. The construct conferring Blasticidin resistance will be introduced to the chromosome using Crispr methodology.

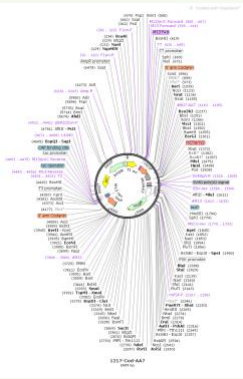


Figure 1. Gene map representation of vector #1217.

In order to examine if the transformation was successful, we ran the DNA from eight bacterial colony samples using gel electrophoresis. Gel

electrophoresis is an experimental method by which DNA fragments are separated according to size. The results are demonstrated below in Figure 2. A band appears around 1,040 kDa in samples Vector, 3, 5, and 7 which is representative of the NheI-NheI fragment conferring Hygromycin resistance. This band indicates that these bacterial samples were not successfully transformed to the Blasticidin resistance. In all samples, a band present at 560 kDa is representative of the fragment between HindIII and HindIII. However, an additional band appeared for samples 2, 4, 6, and 8 around 490 kDa. Since it was difficult to distinguish between the band at 560 kDa and 490 kDa, we ran another gel with samples Vector, 2, 3, 4, 6, and 8. The results are demonstrated in Figure 3, where a band at 560 kDa is visible in all samples. Furthermore, there is a clear appearance of a band at 490 kDa in samples 2, 4, 6, and 8. Since the original NheI-NheI Hygromycin resistance sequence was 1,040 kDa long and was replaced with an NheI-NheI fragment for Blasticidin resistance that was 420 kDa long, the final sequence between ClaI and HindIII should appear around 490 kDa. The presence of this band identifies the samples which have been successfully transfected with Blasticidin resistance.

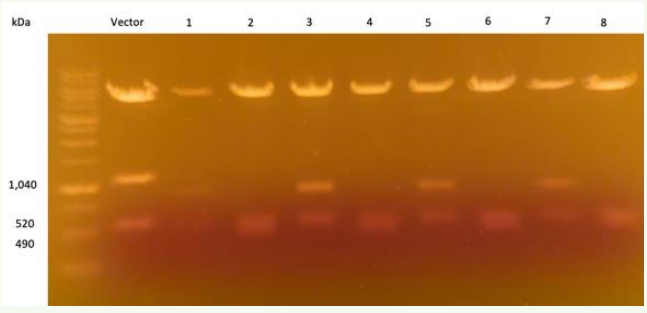


Figure 2. Transformation results examined using gel electrophoresis to compare vector #1217 and eight bacterial samples.

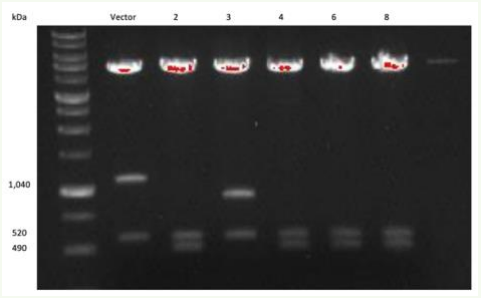


Figure 3. Transformation results repeated using gel electrophoresis. Bacterial samples 2, 4, 6, and 8 demonstrate successful transformation with Blasticidin resistance and, therefore, the degron.

The transformation described above is planned to be used in several mammalian cell types including HeLa, C12, U2OS, and SH-SY5Y cells. In order to study the effects of the transformation on Codanin-1, Auxin was first introduced to transformed HeLa cells. Interestingly, the knockout of Codanin-1 by the degron did not result in cell death or any changes in the cell phenotype. This led to the theory that Codanin-1 is not required in all cell types. Thus, further studies in this area point to the knockout of Codanin-1 in cells in which Codanin-1 is believed to be of significance, such as U2OS or SHSY cells. The effects of Codanin-1 elimination on cellular function following introduction of Auxin in successfully transformed U2OS or SHSY cells has yet to be confirmed.

Previous studies we have conducted have worked towards determining the correct

concentration of Hygromycin to use in order to achieve resistance. However, these studies have been unsuccessful as the concentration used have been found to be lethal. Further study of the optimal concentration of Hygromycin should be performed.