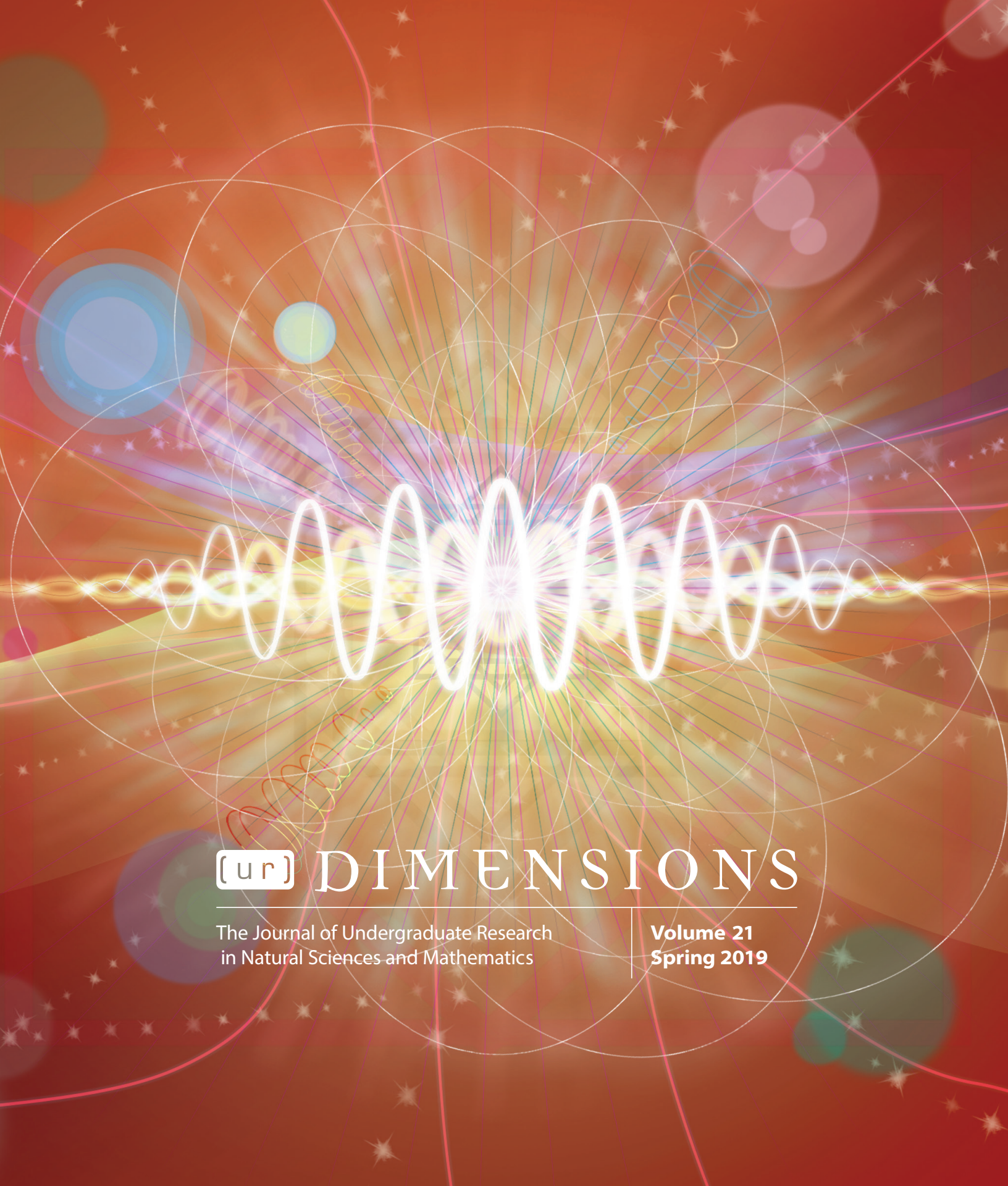


[ur] DIMENSIONS

The Journal of Undergraduate Research in Natural Sciences and Mathematics

Volume 21
Spring 2019

[ur] DIMENSIONS
The Journal of Undergraduate Research in Natural Sciences and Mathematics
Volume **21** | Spring **2019**

Marks of a CSUF graduate from the College of Natural Sciences and Mathematics

GRADUATES FROM THE COLLEGE OF NATURAL SCIENCES AND MATHEMATICS:

Understand the basic concepts and principles of science and mathematics.

Are experienced in working collectively and collaborating to solve problems.

Communicate both orally and in writing with clarity, precision, and confidence.

Are adept at using computers to do word processing, prepare spreadsheets and graphs, and use presentation software.

Possess skills in information retrieval using library resources and the internet.

Have extensive laboratory, workshop, and field experience where they utilize the scientific method to ask questions, formulate hypotheses, design and conduct experiments, and analyze data.

Appreciate diverse cultures as a result of working side by side with many people in collaborative efforts in the classroom, laboratory, and on research projects.

Have had the opportunity to work individually with faculty members in conducting research and independent projects, often leading to the generation of original data and contributing to the research knowledge base.

Are capable of working with modern equipment, instrumentation, and techniques.

DIMENSIONS

DIMENSIONS: The Journal of Undergraduate Research in Natural Sciences and Mathematics is an official publication of California State University, Fullerton. DIMENSIONS is published annually by CSUF, 800 N. State College Blvd., Fullerton, CA 92834.

ABOUT THE COVER

The ongoing exploration using particle collisions has taken our understanding of the physical world into a new era. With each discovery we make, we obtain new ideas, and in turn, drives our sense of learning. We depict an illustration of a perceived effect of colliding with the quantum world with the hopes of motivating the scholars published in this journal to continue to grow their curiousity and contribute in the sciences that shape our world.

DIMENSIONS Editorial Staff and Thanks

EDITOR-IN-CHIEF

Mikayla Mays - Physics

EDITORS

Jessica Sherman - Biological Science

Mayra Silva - Biological Science

Vianney Martinez - Chemistry and Biochemistry

Lindsey Langer - Geological Sciences

Sonali Vyas - Mathematics

Diana Shirley - Physics

ADVISOR

Dr. Colleen McDonough - Assistant Dean for Student Affairs

GRAPHIC DESIGN

Gabriel Martinez - Layout Editor and Cover Designer

COLLEGE OF NATURAL SCIENCES & MATHEMATICS

Dr. Marie Johnson - Dean

Dr. Sean E. Walker - Associate Dean

Dr. Merri Lynn Casem - Chair, Department of Biological Science

Dr. Peter de Lijser - Chair, Department of Chemistry and Biochemistry

Dr. Jeffrey Knott - Chair, Department of Geological Sciences

Dr. Alfonso Agnew - Chair, Department of Mathematics

Dr. Ionel Tifrea - Chair, Department of Physics

Special Thanks To

The CNSM administrative staff for their support, the faculty for advising students and providing them with valuable research opportunities, and Dr. Colleen McDonough for her guidance and oversight.

Table of Contents

54 Geological Sciences

54 Using U-Pb Dating to Determine the Maximum Depositional Ages and Provenance of the Cow Creek and Fairview Metasedimentary Pendants, southern Sierra Nevada Mountains
- **Christian Concha**
- **Advisor: Diane Clemens-Knott, Ph.D.**

61 Conditions And Persistence Of Mcdonald Well, Bird Spring, And North Bird Spring In The Mojave Desert, San Bernardino County, California
- **Eddie Reyes**
- **Advisors: Jeffrey R. Knott, Ph.D., and Anna L. Garcia**

62 Extending our view of a shallow-dipping Sierra Nevada Fault System south of Mount Whitney near Lone Pine, California
- **Jackson Flanagan**
- **Advisor: Phillip A. Armstrong, Ph.D.**

65 The Nature of Textural and Compositional Variability in the Lake Vernon Granodiorite in Western Yosemite National Park, CA.
- **Jamie Hayward and Cullen Scheland**
- **Advisor: Valbone Memeti, Ph.D.**

66 Using Holocene sediments from North Yolla Bolly Lake in the Northern Coast Range (CA) to investigate the California precipitation dipole
- **Judith A. Avila, Joseph A. Carlin, Glen M. Macdonald, Alexandre Woodward, Jenifer Leidelmeijer, Jiwoo Han, and Benjamin Nauman**
- **Advisor: Matthew E. Kirby, Ph.D.**

67 Clast Provenance And Detrital Zircon Dating Of Pliocene Fluvial Conglomerate: Implications For Paleohydrology Of Crooked Creek And Deep Springs Valley, California
- **Lindsey M. Langer**
- **Advisors: Jeffrey R. Knott, Ph.D., Diane Clemens-Knott, Ph.D., and Anna Garcia**

68 Mathematics

68 Incorporating Migration into an Age-Structured Population Model with Variable Mortality Rate
- **Gabriel Martinez and Freddy Nungaray**
- **Adviros: Laura Smith Chowdhury, Ph.D.**

79 Ivertible Polynomials and their Mirrors
- **Kenneth Eaves**
- **Advisor: Christopher Lyons, Ph.D.**

89 Exploring Multi-Colored Peg Solitaire
- **Melissa Wong, Sonali Vyas, and Alexxis De Lamere**
- **Advisors: Roberto Soto, Ph.D., and Tara Davis, Ph.D.**

90 Longevity and Immune Defense, Evolutionary Trade-Offs in *Drosophila melanogaster*
- **Nayelie Y. Benitez Santos and Han Yin**
- **Advisors: Parvin Shahrestani, Ph.D., and Sam Behseta, Ph.D.**

99 Gone With The Flu
- **Kristina Rosete, Taylor Salgado, Sonali Vyas**
- **Advisor: Roberto Soto, Ph.D.**

100 Economical Generating Sets of the Monoid of Patial Order-preserving Injections
- **Saul Lopez**
- **Advisor: Scott Annin, Ph.D.**

114 Physics

114 Student Blending in Math and Physics Integration Problems
- **Anthony Pina**
- **Advisor: Michael Loverude, Ph.D.**

115 The Stochastic Force Spectrum of a Micro-Swimmer
- **Corbyn Jones and Mauricio Gomez**
- **Advisor: Wylie Ahmed, Ph.D.**

Diffusible Iodine-Based Contrast-Enhanced Computed Tomography to Study Surf Perch (*Embiotocidae*) Reproductive Anatomy

Andrew Jaramillo
Advisor: Kristy L. Forsgren, Ph.D.
Department of Biological Science, California State University, Fullerton

The surfperches (Embiotocidae) are a family of fishes that reside along the Pacific coast of California. A distinctive life history trait of the surfperches is that they exhibit internal fertilization, also known as viviparity. Currently, it is unclear what tissues and/or organs are directly involved in surfperch sperm transfer from the male into the female reproductive tract. We are am interested in the natural (i.e., not dissected) relationship of the testes, genital papilla, and paired flask organs of the anal fin. Our study focused on the comparative reproductive anatomy of male surfperches: walleye surfperch (*Hyperprosopon argenteum*), black perch (Embiotoca jacksoni), shiner perch (*Cymatogaster aggregata*), and dwarf perch (*Micrometrus minimus*). Our study objective was to use a common medical diagnostic technique, computed tomography (CT) scanning, to gain a better understanding of surfperch reproduction. Fish were collected via beach seine, preserved in formalin, and stained with Lugol's iodine solution prior to CT scanning [diffusible iodine-based contrast-enhanced computed tomography (DICECT)]. This unique technique capitalizes on iodine's properites of hardening soft tissues so that they are visualized in contrast with hard (i.e., bone) tissue on the scan. Digital files (dicom) from high resolution DICECT scans were segmented by color-coding different tissues using the computer program ITK-SNAP in order to label structures and create a 3D image (Fig. 1). Each anatomical structure of interest was coded with a different color but the same color for each structure among different species in order to directly compare species (e.g., testes are blue, urinary bladder is orange). Our emerging research will provide additional insight into the functional morphology of surfperch reproduction and increase of understanding of internal fertilziation of marine fishes.

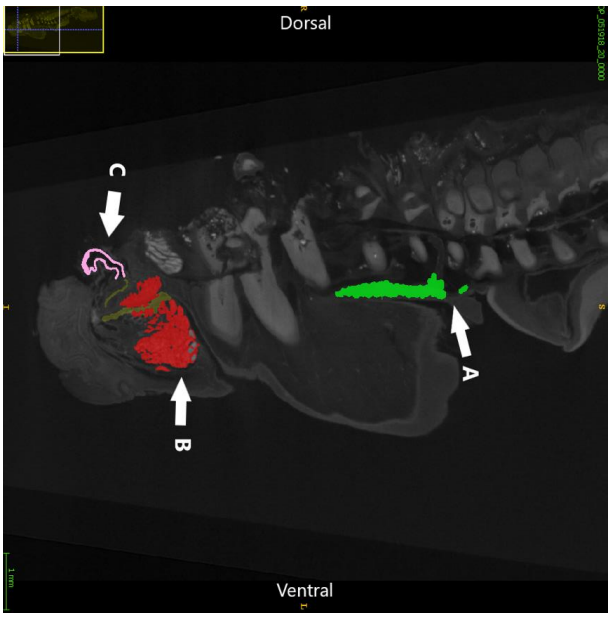


Figure 1. *Dwarf Perch (Micrometrus minimus) anal fin cross section of a computed topography (CT) scan indicating the relationship between the flask organ (A), genital papilla (B), and the sperm duct (C).*

Effects of Fire on a Blackbrush and Joshua Tree Woodland in the Mojave Desert

Ariana M. Sanchez, Kassandra M. Rodriguez, Syndee R. Dunn, Andrew M. Jaramillo, and M. Kim Koval
Advisors: Jennifer L. Burnaford, Ph.D., and Darren R. Sandquist, Ph.D.

Department of Biological Science, California State University, Fullerton

ABSTRACT

Desert organisms are often highly adapted to extreme conditions, however, native desert communities can be slow to recover after major disturbances (e.g., wildfire) and are prone to invasion by non-native grasses. We examined plant community composition of a Mojave Desert Joshua tree woodland recovering from a 2005 fire by comparing adjacent burned and unburned areas in 2018. We hypothesized that fire disturbance would increase abundance of grasses and annuals relative to an unburned area but decrease perennial species richness and diversity. We quantified the plant community using line transects, 50 m^2 belt transects, and 0.25 m^2 quadrats. Cover of non-native grasses did not differ between burned and unburned areas, but percent cover of native grasses was approximately ten times higher in the burned area than the unburned area. There was an average of 5.7 ± 0.6 (mean ± SE) perennial species in the burned area and 9.0 ± 0.7 in the unburned area. The diversity of perennial plants was higher in the unburned area (mean H'= 1.742) than in the burned area (mean H' = 1.075). We found an average density of 1.94 ± 0.1 individuals per m^2 for annual plants in the burned area, but none were found in the unburned area. Despite 13 years of recovery, we found that the burned area community composition and cover remained very different from the unburned area, but that active management may not be necessary to prevent conversion to a non-native grass system.

Enhancing reference genomes for cross-cultivar functional genomics in apples

[1]Brittany Cook, [1]Mathew C. Simenc, and [2]Heidi Hargarten
Advisors: [1]Joshua P. Der, Ph.D., and [2]Loren Honaas, Ph.D.
[1]*Department of Biological Science, California State University, Fullerton*
[2]*USDA Agricultural Research Service, Tree Fruit Research Lab, Wenachee, WA*

ABSTRACT

Apples are one of the most valuable tree fruit crops in the United States, and modern advances in controlled-atmosphere storage has made apples commercially available year-round. However, a number of economically important cultivars are susceptible to a number of storage disorders impacting post-harvest tree fruit quality. Understanding the physiology of these disorders has become an important component of apple breeding programs. While the apple genome was sequenced in 2010, the reference genome comes from a double haploid version of the Golden Delicious cultivar, but substantial genetic diversity among apple cultivars has represented a barrier to using genomic approaches in other economically relevant apple cultivars. We are using an RNA-seq and whole genome resequencing approach to characterize genetic diversity within apples and to improve the utility of the reference genome for cross-cultivar functional genomics. We have obtained and analyzed sequencing data from 62 samples of Honeycrisp (HC) and Granny Smith (GS) representing both healthy and diseased apples undergoing a number of storage treatments to identify genes related to post-harvest fruit quality. We selected 29 candidate genes, 15 from GS and 14 from HC, to explore in detail the genetic differences between cultivars of interest and the reference cultivar, Golden Delicious. Gene expression levels for these genes were also validated using qPCR. We hypothesis that correlation between genes of interest and hotspots of variation emphasizes the need to explore these differences in greater detail to move forward in applied, functional tree-fruit genomics. With a subset of transcriptome samples from GS, we evaluated the mapping rate of RNA-seq reads to the Golden Delicious reference. Reads were aligned to the reference using HISTAT2 and the transcript models were determined using StringTie. We found that only 70% of our apple transcriptome data confidently mapped to the reference genome, highlighting the need for an improved reference. We are building de novo transcript models using Trinity and have re-sequenced the GS genome using both Illumina and PacBio whole genome shotgun reads. We are in the process of assembling the GS genome and quantify the distribution of genetic variation between GS and Golden Delicious apples.

What is the relationship between longevity and innate immunity in *Drosophila melanogaster*?

Elnaz Bagheri and Melanie Martinez
Advisor: Parvin Shahrestani, Ph.D.
Department of Biological Science, California State University, Fullerton

ABSTRACT

Bacteria, viruses and fungi create infectious diseases that cause morbidity and mortality. This research focuses on two questions using fruit flies, *Drosophila melanogaster*, as the study system. First, what is the relationship between longevity and immune defense in short-lived and long-lived *Drosophila* populations? Second, what is the effect of chronological age on immune defense in the short and long lived populations? *D. melanogaster* populations were experimentally evolved to become differentiated in aging rates and longevity. Then five replicate short-lived populations were compared to five replicate long-lived populations for differences in immune defense against the fungal pathogen Beauveria bassiana. Immune defense was measured at multiple ages throughout the lifespan for all populations. Based on our preliminary results, post-infection survival is lower in short-lived populations at all chronological ages compared to the same ages in long-lived populations, and immune defense declines with age. After the age-specific immune defense phenotype is fully characterized, we will look for candidate immunesenescence genes using the existing whole-genome sequencing data of these ten populations. Results from this study can potentially inform the use of B. bassiana in biological control against mosquito vectors of malaria and dengue, which are genetically similar to *D. melanogaster*. Moreover, due to similarities between the innate immunities of fruit flies and humans, results from this study can inform the need for personalized medicine, such as treating old and young patients differently.

Size Matters: Comparative Morphology of Rockfish Urogenital Papilla

Holly L. Suther
Advisor: Kristy L. Forsgren, Ph.D.
Department of Biological Science, California State University, Fullerton

ABSTRACT

Rockfishes are a popular recreational and commercial fishery in California. Rockfish catches account for 20% of the state's commercial harvest, which contributes $550 million dollars annually to California's economy. As an economically and ecologically important fishery, regulatory agencies monitor the status of rockfish populations. However, there is no accurate method to identify sex, data that has important implications for fishery management. The objective of our study is to increase our understanding of rockfish reproduction by characterizing the male urogenital papilla of various rockfish species in order to establish a reliable method of externally identifying sex in the field. Rockfishes [starry *(Sebastes constellatus)*, vermilion (Sebastus miniatus), squarespot *(Sebastes hopkinsi)*, blue rockfishes *(Sebastes mystinus)*] were collected in southern California via hook and line. Gonadal tissues and genital papilla were dissected and preserved, then embedded in paraffin wax. Tissues were sectioned using a rotary microtome, stained, and histologically examined. Additional specimens were borrowed from the Natural History Museum of Los Angeles County for external measurements. We have determined that the morphology of the urogenital papilla is species-specific and can be used to accurately identify males. Our future work includes describing the morphology of additional rockfish species to develop a comprehensive tool for fishermen and agency biologists.

Do Rural and Suburban Coyotes Have Different Diets? A Comparison Using Stable Isotope Analysis

Kaela Manalo
Advisor: Paul Stapp, Ph.D.
Department of Biological Science, California State University, Fullerton

ABSTRACT

Coyotes are opportunistic carnivores that thrive in urban and suburban settings because of their tolerance of humans and their ability to take advantage of both natural and anthropogenic foods. Many previous studies of coyote diets have used conventional methods (scat or stomach contents analysis), which provides a measure of diet over one or two meals and may be biased against certain prey types. We used stable carbon and nitrogen isotope analysis to compare the longer-term, integrated diets of coyotes from two rural areas of California (Fresno and Mendocino Counties; N = 21) to diets of coyotes from suburban southern California (Los Angeles and Orange Counties; N = 126). We analyzed muscle tissue from the lower jaws of coyotes killed by vehicles or by animal control agencies. For a subset of 15 rural animals, we also analyzed samples of guard hairs and whiskers to determine if the isotope values from tissues that could be collected non-destructively were predictably similar to those of muscle tissue, which must be taken from a dead animal. Muscle tissue of rural coyotes was significantly depleted in both δ 13 *C* and δ 15 N compared to suburban coyotes, and muscle δ 13 *C* of coyotes from Mendocino County was also significantly lower than that of ones from Fresno County. Muscle δ 15 N values did not differ between rural sites, suggesting that these coyotes occupied the same trophic level.Muscle δ 13 *C* was significantly depleted compared to δ 13 *C* of both whisker and hair tissue from the same animal, but whisker and hair tissue did not differ from one another. In contrast, δ 15 N of muscle tissue was significantly enriched compared to δ 15 N of whiskers and hair, which were only marginally different from one another. Our results suggest that coyotes from rural areas have a very different diet, likely composed primarily of herbivorous mammals like rabbits, rodents, and possibly livestock, than their suburban counterparts, whose food habits were highly variable. The consistent relationships between isotope values for muscle and whisker samples indicate that whiskers may be a suitable, non-destructive source of tissue to determine long-term diet using stable isotope analysis.

Experimental Evolution for Longevity Differentiation in *Drosophila melanogaster*

Karen Walsh and Melanie Garcia
Advisor: Parvin Shahrestani, Ph.D.
Department of Biological Science, California State University, Fullerton

ABSTRACT

The evolutionary theory of aging proposes that the forces of natural selection start to decline after the first age of reproduction in a population, and continue to steadily decline until the last age of reproduction, where these forces stabilize at zero or become weaker than random genetic drift. Past studies have used *Drosophila melanogaster* populations to show that gradually postponing the first age of reproduction, postpones the age at which the forces of natural selection begin to drop, and results in delayed aging and increased longevity. However, few studies have focused on identifying the differences between short-lived and long-lived populations of D. melanogaster on a genomic scale. The first genome-wide analysis of an experimental evolution study in a sexual organism found significant differences in single nucleotide polymorphisms (SNPs) between D. melanogaster populations selected for accelerated development and their controls, but longevity was not the main focus. Genomic studies involving longevity remain an under researched topic in experimentally evolved populations. In this study, we aim to identify candidate genes involved in longevity using a genome-wide analysis of experimentally evolved short-lived and long-lived populations of *D. melanogaster* . Long-lived populations are created by progressively postponing the first age of reproduction from 14 days to 70 days. To maximize the statistical power of finding these candidate genes, we use ten replicate populations selected for increased longevity, maintained on 70-day discrete generation cycles, and ten replicate populations that are treated as matched controls, maintained on 14-day discrete generation cycles. We have completed six generations of laboratory adaptation followed by four generations of selection for postponed reproduction, and will continue this process for at least 20 generations. Samples of the long-lived and short-lived flies are taken at various generational time points throughout the duration of the experiment and are preserved for genomic sequencing and comparison. Identifying regions of the genome that are differentiated between short and long-lived populations may provide candidate genes for human longevity. Determining the trajectories of adaptive alleles throughout experimental evolution could provide conservation biologists with tools to make predictions about how species may evolve in changing wild environments.

The Effect of Pollutant Exposure on Fathead Minnow Nuptial Tubercle Development

Zsuzsanna Papp and [2]Alvina Mehinto
Advisors: [1]Kristy Forsgren, Ph.D.
[1]Department of Biological Science, California State University, Fullerton
[2]Southern California Coastal Water

ABSTRACT

The fathead minnow (*Pimephales promelas*) is an ideal model for examining the effects of environmental contaminants and has been used previously to establish assays for ecotoxicological impacts. Male fathead minnows develop a fatpad (i.e., grey spongy growth on their nape and back) and nuptial tubercles (i.e., numerous white bumps around the snout), which are distinctive secondary sexual characteristics in sexually mature fish. *Our hypotheses are that: 1) nuptial tubercle characteristics are associated with reproductive developmental stages of sexually maturing males; and 2) exposure to potentially polluted waters of the Los Angeles River will influence nuptial tubercle formation*. Fathead minnows were exposed to wastewater effluent near the Tillman Water and LA-Glendale Reclamation Plants for three weeks during May – June 2018. Fish cages which consisted of two identical tanks (i.e., A and B) were setup with 8 fish per tank and a total of 16 fish in each cage (n = 4 cages). Fish cages were placed near the shoreline and the water was circulated into the cages via pumps. Laboratory control treatments were set up at the same time; a negative control with filtered, "clean" seawater and a positive control with the hormone, estrone, added to filtered seawater (100 ng/L). On day 0, 13 fish were terminally samples to establish baseline measurements for the project. After the exposure, fish were euthanized and the length (mm), weight (g), and gonadal weight (g) were recorded. Macrophotography documenting the head, particularly the snout with nuptial tubercles, were taken. In order to examine the nuptial tubercles, the front portion of the head was fixed using 10% neutral-buffered formalin for 48 hours then transferred to 70% ethanol until histological processing. The head region was first treated with decalcifying solution, then processed in a series of graded ethanols and xylene and infiltrated and embedded in paraffin wax. Tissues were sectioned in the coronal direction using a rotary microtome to a thickness of 5µm, mounted on slides, and stained with hematoxylin and eosin or Mallory Trichrome. Using Image J, photographs will be used to measure the area of the tubercles and document the pattern of nuptial tubercles. Tissue sections will be examined under light microscope to characterize of the nuptial tubercles by counting the number of tubercles and measuring the length (mm) and height (mm) of each tubercle. The data collection and analysis are ongoing. To date, we have determined that the chromatophore layer appears to reach into the nuptial tubercles from the layer within the dermis. We aim to provide a detailed histological description of nuptial tubercles to be used by a variety of biologists seeking to understand the impact of pollution on reproduction and exposure-based fathead minnow studies, which have primarily focused on pathology of gonadal development. Nuptial tubercle characterization will allow for external assessment of fathead minnow exposure experiments without the need for invasive dissection techniques. Overall, our research is important in order to demonstrate the impact of urban runoff, which could affect the sexual development of organisms exposed to effluent.

Comparison of Alpha-2-Macroglobulins from Swine and Humans and their Copper Binding

[1]Alejandra Garcia, [1]Cameron Marzvaan, and [2]Enrico Falcone
Advisors: [1]Maria C. Linder, Ph.D., and [2]Peter Faller, Ph.D.

[1]*Department of Chemistry and Biochemistry, California State University*

[2]*Department of Chemistry, University of Strasbourg, France*

Alpha-2-macroglobulin (a2M) is a member of the macroglobulin family of blood plasma proteins. It has a variety of functions, ranging from trapping proteases and transporting zinc to binding and transporting inflammatory and anti-inflammatory cytokines and, growth factors and the small peptide regulating iron metabolism (hepcidin) [1]. A2M is the main form found in most mammals, while the alpha1-inhibitor3 form dominates in the plasma of rodents. Some years ago, we determined in rats that a large plasma protein we named transcuprein was involved in transporting copper to the liver (exchanging copper with albumin), and upon purification found it was alpha1-inhibitor3[2]. We then showed that human a2M also bound Cu(II) tightly, and that this copper was readily delivered by a2M to cultured human cells[3]. As we have been studying aspects of copper metabolism in pigs, we decided to also investigate the structure and copper binding of pig a2M to compare it to that of human a2M. Human a2M was purified from the heparinized plasma of human volunteers (under an approved university IRB protocol), using the established procedure for human a2M that uses a combination of PEG 8000 fractionation and Zn(II)-immobilized metal affinity chromatography. The resulting samples were separated in large pore size exclusion chromatography on Sephacryl S300. Human a2M eluted had one large peak eluting with an MW of ~900 kDa, and a much smaller peak of about 400 kDa. By SDS-PAGE, both peaks showed the expected 180 kDa band, indicating that most of the human a2M was a tetramer, with a very small portion as a dimer. The copper content of the human a2M was 4 Cu atoms per tetramer (not taking into account potential non-specific binding), but was 2 Cu atoms per tetramer when copper was first removed and added back in the presence of human or rat albumin. The same purification procedure, this time beginning with heparinized Yorkshire pig plasma, eluted a protein about half the size (~360 kDa) of the human a2m, indicating a dimer. Non-denaturing PAGE gave single bands and SDS-PAGE revealed a single band subunit with size of ~70 kDa. The Cu content of the pig protein was 4 Cu atoms per dimer. Preliminary data obtained by EPR gave similar but not identical profiles for the two proteins and indicated there were two binding sites, the main one with Cu(II) atoms most likely bound to three Ns and one O. However, mass spectrometry revealed the pig protein as hemopexin, identifying hemopexin as a copper transporter in swine. We are now conducting further studies on hemopexin as a copper carrying protein and purifying pig a2m.

REFERENCES

[1] M.C. Linder (2016) *Metallomics* DOI: 10.1039/c6mt0010 3c

[2] N.M. Liu et al. (2007) *J. Nutr. Biochem.* 18: 597-608

[3] T.Z. Kidane et al. (2012) *Biometals* 25: 697-709

Photoinduced Oxidative Cyclization Reactions with Thiophene and Furan Rings as Nucleophiles

Crystal Paulus, Catherine Taylor and Abdullah Alshreimi
Advisor: Peter de Lijser, Ph.D.
Department of Chemistry and Biochemistry, California State University, Fullerton

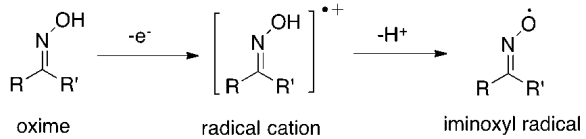
ABSTRACT

Oxime and oxime ethers form reactive intermediates, such as radicals and radical cations, when exposed to oxidative conditions. Previous studies have shown that in these oxidative processes, oximes tend to form an iminoxyl radical and oxime ethers tend to form a radical cation intermediate. These reactive intermediates can be used to synthesize heteroaromatic structures via intramolecular reactions with built-in nucleophiles or radical traps. Previous work has demonstrated that oxime ethers with built in aromatic groups can form a cyclized product (phenanthridine) when the aromatic ring behaves as a nucleophile attacking the radical cation center (Nitrogen atom) of the oxime ether. To further evaluate the general applicability of these types of reactions, we have extended our studies to built-in nucleophiles that contain heteroatoms, including pyridine, thiophene and furan rings. Photolysis of the oxime ethers with built-in nucleophiles in the presence of a photosensitizer (chloranil (CA) or 9,10-dicyanoanthracene (DCA)) resulted in the formation of the desired products, albeit in low yields for most of the substrates. The thiophenyl derivatives of the oxime ethers gave the most promising results, followed by the furans. The 2-thiophene derivative reacted best in the presence of DCA at 420 nm, whereas the 3-thiophene derivative reacted best in the presence of CA at 350 nm. It is unclear at the moment what causes this different behavior.

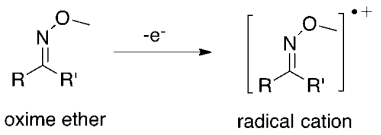
INTRODUCTION

Oxime and oxime ethers are widely used as drugs, additives, and pesticides.[1] Drugs or pesticides that contain oximes or oxime ethers will be recognized in the human body as xenobiotic, meaning foreign molecules, which are going to be exposed to oxidative processes via metabolism by enzymes such as Cytochrome P450.[2] When oximes and oxime ethers are subjected to oxidative processes, reactive intermediates can be formed, such as radicals and radical cations,[3-5] which can cause damage to DNA.[6] These reactive intermediates can potentially also be useful in synthetic routes for preparing novel molecules. In order to take advantage of this potential reactivity, a better understanding of their reactivity is needed. To study these reactive intermediates in the laboratory we use photoinduced electron transfer (PET) to generate them under controlled conditions so that we can focus on understanding the structure and reactivity of these radicals and radical ions.

Previous studies have suggested that under oxidative conditions, oximes tend to form an iminoxyl radical (Scheme 1) and oxime ethers tend to form a radical cation (Scheme 2) intermediate. The oxime radical cation is short-lived because it is very acidic and deprotonates rapidly, forming the iminoxyl radical, which is fairly stable and relatively unreactive. Trapping



Scheme 1. *Oximes form iminoxyl radical in a two-step process with a presence of a photosensitizer.*



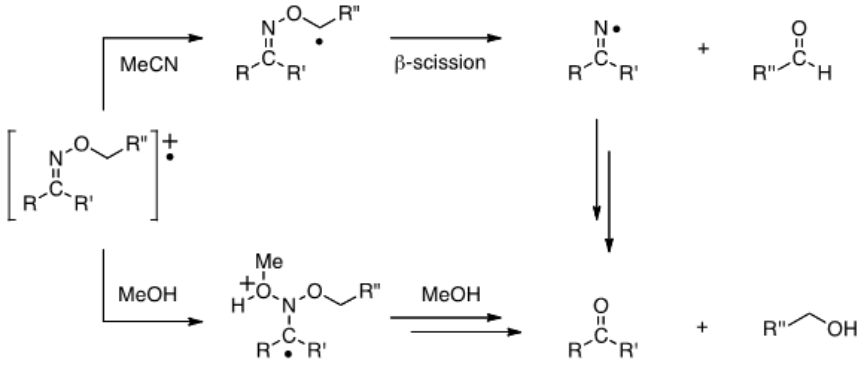
Scheme 2. *Oxime Ethers form a radical cation in a one-step process with a presence of a photosensitizer.*

the oxime radical cation has been very difficult, making oximes less reactive in PET reactions.

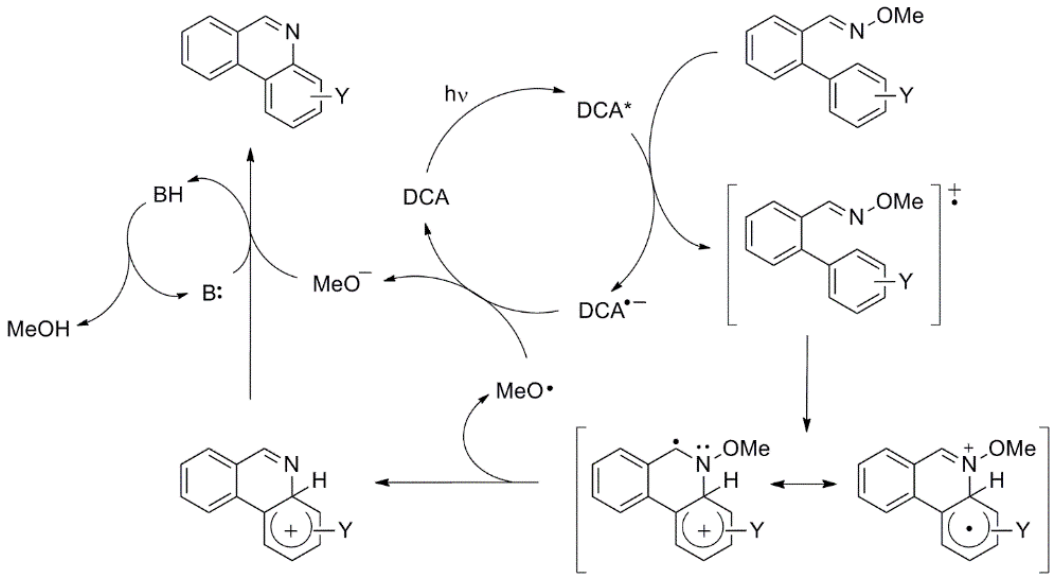
The PET reactions of oxime ethers have been shown to be solvent-dependent. In a non-nucleophilic solvent (acetonitrile) the product formation can be explained by a mechanism that involves electron transfer followed by proton transfer (α to the oxygen) and subsequent *β*-scission (Scheme 2; top pathway). In a nucleophilic solvent (methanol) the mechanism is most consistent with a sequence involving electron transfer, followed by a nucleophilic attack on the nitrogen, a methanol-assisted [1,3]-proton transfer, and subsequent loss of an alcohol (Scheme 3, bottom pathway).[7] In a non-nucleophilic solvent the product formation can be explained by a mechanism that involves electron transfer followed by proton transfer (α to the oxygen) and subsequent *β*-scission (Scheme 3; top pathway). In a nucleophilic solvent the mechanism is most consistent with a sequence involving electron transfer, followed by a nucleophilic attack on the nitrogen, a methanol-assisted [1,3]-proton transfer, and subsequent loss of an alcohol (Scheme 3, bottom pathway).[7]

The PET reactions of oxime ethers have been shown to be solvent-dependent. In a non-nucleophilic solvent (acetonitrile) the product formation can be explained by a mechanism that involves electron transfer followed by proton transfer (α to the oxygen) and subsequent *β*-scission (Scheme 2; top pathway). In a nucleophilic solvent (methanol) the mechanism is most consistent with a sequence involving electron transfer, followed by a nucleophilic attack on the nitrogen, a methanol-assisted [1,3]-proton transfer, and subsequent loss of an alcohol (Scheme 3, bottom pathway).[7] In a non-nucleophilic solvent the product formation can be explained by a mechanism that involves electron transfer followed by proton transfer (α to the oxygen) and subsequent *β*-scission (Scheme 3; top pathway). In a nucleophilic solvent the mechanism is most consistent with a sequence involving electron transfer, followed by a nucleophilic attack on the nitrogen, a methanol-assisted [1,3]-proton transfer, and subsequent loss of an alcohol (Scheme 3, bottom pathway).[7]

The behavior of oxime ether radical cations in the presence of nucleophiles led us to investigate the reactions of oxime ethers with built-in nucleophiles to see if intramolecular reactions would occur. A series of 2'-arylbenzaldehyde oxime ethers were synthesized and shown to generate the corresponding phenanthridines upon irradiation (420 nm) in the presence of 9,10-dicyanoanthracene (DCA) in acetonitrile. Mechanistic studies suggest that the oxidative cyclization reaction sequence is initiated by an electron transfer step followed by nucleophilic attack of the aryl ring onto the nitrogen of the oxime ether (Scheme 4).[8]



Scheme 3. *Mechanistic pathways of oxime ethers upon one-electron oxidation.*[7]



Scheme 4. *Proposed Mechanism for 2-arylbenzaldehyde O-methyl oxime ethers.8*

2'arylbenzaldhyde oximes did not undergo cyclization; the major product in those reactions was the parent aldehyde.

To further evaluate the general applicability of these types of reactions, we have extended our studies to built-in nucleophiles that contain heteroatoms, including pyridine, thiophene and furan rings, shown in Figure 1 below.

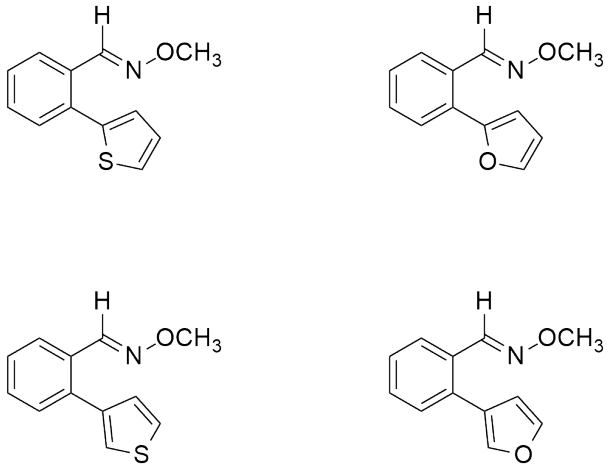
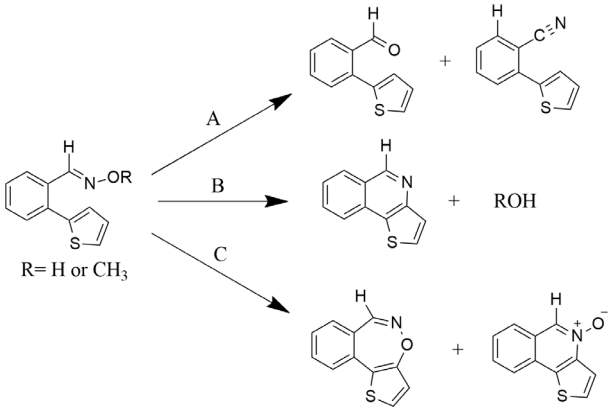


Figure 1. *Oxime Ethers with different heteroaromatic groups that were synthesized and studied.*

Under these photooxidative conditions, oximes and oxime ethers are hypothesized to have three major routes of reactivity (Scheme 5). Pathway A proposes to form the parent aldehyde and nitrile compound which is expected to be the dominant pathway under certain conditions. Quinones (CA) tend to favor the formation of the nitrile. The products from this pathway are easily recognized in ^{1}H NMR spectrum and can be verified by comparison of the reaction mixture to the authentic compound. Pathway B shows the targeted formation for the cyclized product "phenanthridine derivative." Characteristic information from the NMR spectra will be formation of a new signal representing the iminyl hydrogen; the expected chemical shift will be similar to that of the one observed in previous studies.[8] This type of cyclization is characteristic of oxime ethers and results in the formation of an alcohol as a byproduct depending on the ether group. In this study, O-methyl oximes are used and would produce methanol as a byproduct if they underwent cyclization through pathway B. Methanol can be observed in the ^{1}H NMR spectrum at approximately 3.30 ppm. At the same time the signal of the methoxy group in the starting oxime ether will disappear. Pathway *C* which shows a cyclization pathway that is thought to

involve iminoxyl radical as an intermediate. This would only be expected for oximes because oxime ethers react via their corresponding radical cations. These structures are more difficult to analyze in the NMR spectra although previous studies have generated N-oxides and their NMR spectra are known.[9]



Scheme 5. *A diagram shows the proposed hypothesis for the outcome of photo-oxidation processes upon o-aryl heteroaromatic (2-thiophene) oximes and oxime ethers.*

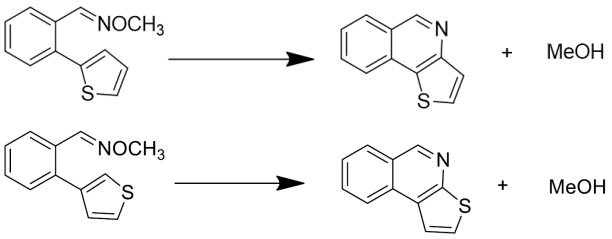
RESULTS AND DISCUSSION

The o-aryl heteroaromatic aldehydes oxime and oxime ether derivatives were prepared according to standard literature methods with minor modifications.[10,11] Our initial photochemical experiments utilized the conditions established in previous studies.[8, 12-15] All reactions were carried out on a small scale in NMR tubes using deuterated acetonitrile in order to follow the reaction progression by ^{1}H NMR. The oxime and oxime ether derivatives we exposed to two different conditions: 350 nm with CA as the sensitizer and 420 nm with DCA as the sensitizer. None of the oxime derivatives showed any sign of cyclization and therefore only the oxime ether derivatives will be considered and discussed.

Thiophene Derivatives

Irradiation of both thiophene derivatives in the presence of a photosensitizer resulted in cyclization, as seen from the appearance of specific signals in the aromatic region. In addition, the formation of methanol is a predominant characteristic of the oxidative cyclization reaction of the O-methyl oxime derivatives:

Figure 2 shows the reactivity of (*E*)-2-(thiophen-



2-yl)benzaldehyde O-methyl oxime when irradiated in the presence of CA (350 nm) or DCA (420 nm) after 8 hours. With CA as the sensitizer, the only result observed was *E*/*Z* isomerization, whereas with DCA, several new signals were observed after 8 hours of irradiation, suggesting formation of a new product.

Figure 3 clearly shows the disappearance of the singlet at 8.2 ppm and the formation of a new singlet at 9.25 ppm, which is consistent with the change in environment for the aldehyde proton in going from the starting oxime to the ring-closed structure.

The product from the reaction was positively identified as thieno[3,2-c]isoquinoline (the proposed cyclic product) when it was isolated and analyzed by NMR (Figure 4).

A similar experiment with (*E*)-2-(thiophen-3-yl) benzaldehyde O-methyl oxime gave different results (Figure 5). Irradiation of (*E*)-2-(thiophen-3-yl) benzaldehyde O-methyl oxime in the presence of CA gave a clean product mixture with new signals appearing at 9.2, 8.4, 8.2, 8.0, 7.9, 7.75, and 7.7 ppm. The mixture of the reaction with DCA on the other hand shows the signals observed in the reaction with CA as well as several other signals, suggesting the formation of multiple products.

Figure 6 clearly shows the disappearance of the singlet at 8.2 ppm and the formation of a new singlet at 9.2 ppm, which is consistent with the change in environment for the aldehyde proton in going from the starting oxime to the ring-closed structure.

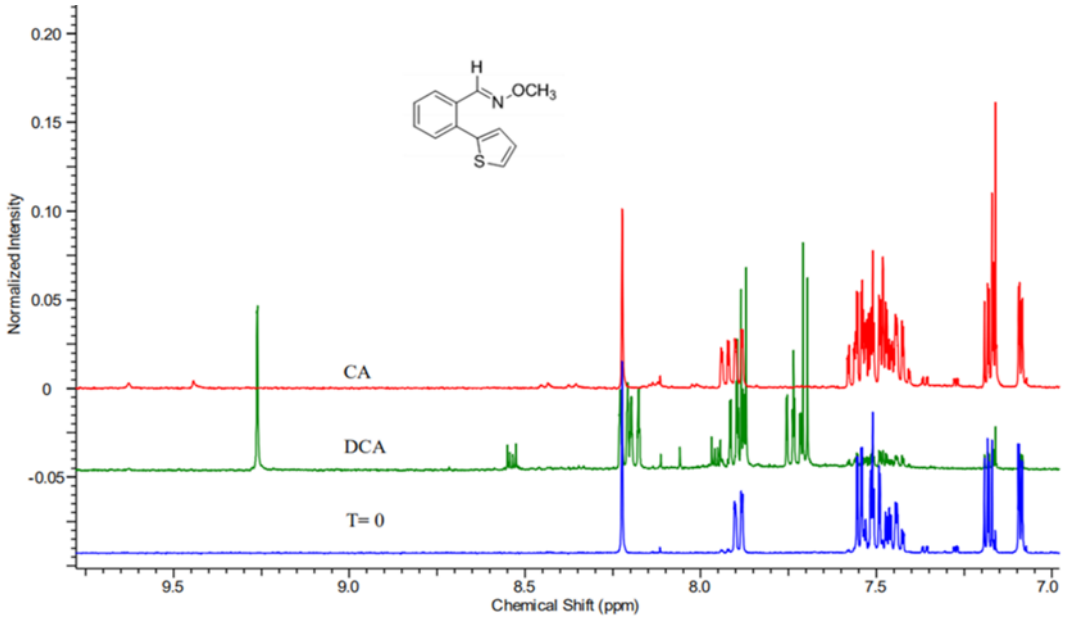


Figure 2. *Overlaid ^{1}H NMR spectra displaying the photolysis reaction of (E)-2-(thiophen-2-yl)benzaldehyde O-methyl oxime at t=0 and t=8 hours for both Chloranil (CA) and 9,10-dicyanoanthracene (DCA).*

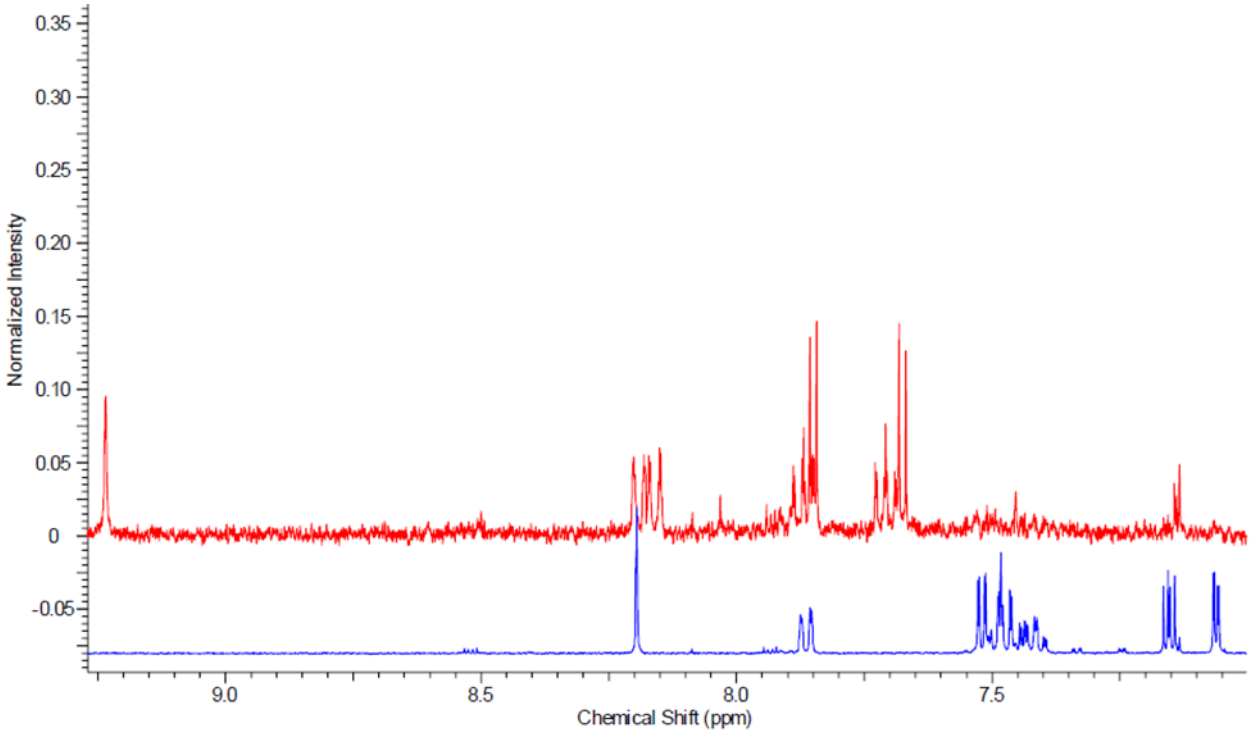


Figure 3. *Overlaid ^{1}H NMR spectra displaying the photolysis reaction of (E)-2-(thiophen-2-yl)benzaldehyde O-methyl oxime at t=0 (blue) and t=8 (red) hours for 9,10-dicyanoanthracene (DCA).*

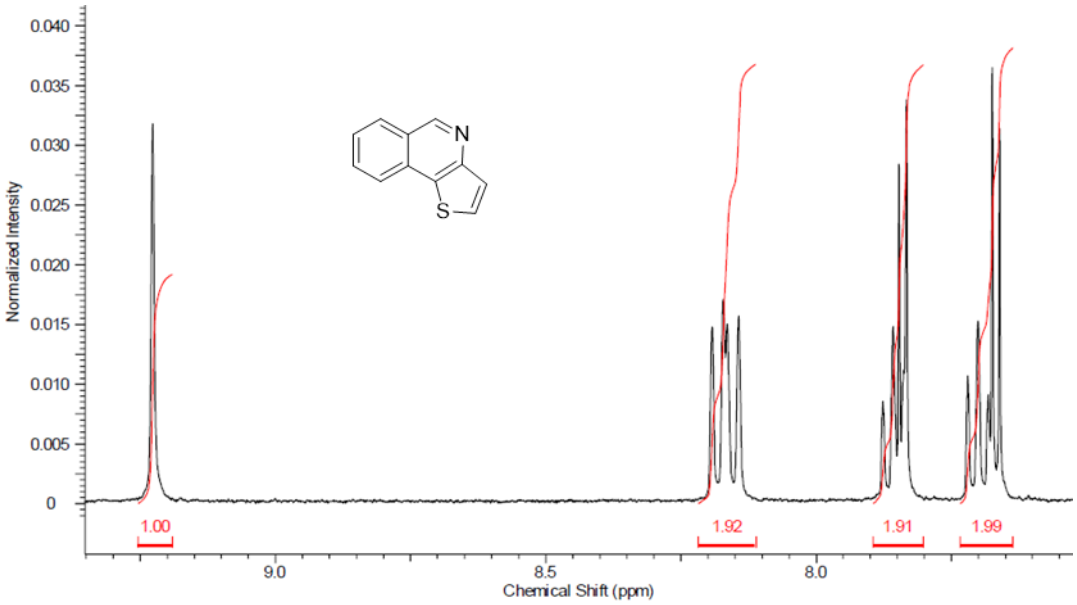


Figure 4. *Isolated product from the reaction of (E)-2-(thiophen-2-yl)benzaldehyde O-methyl oxime with 9,10-dicyanoanthracene (DCA).*

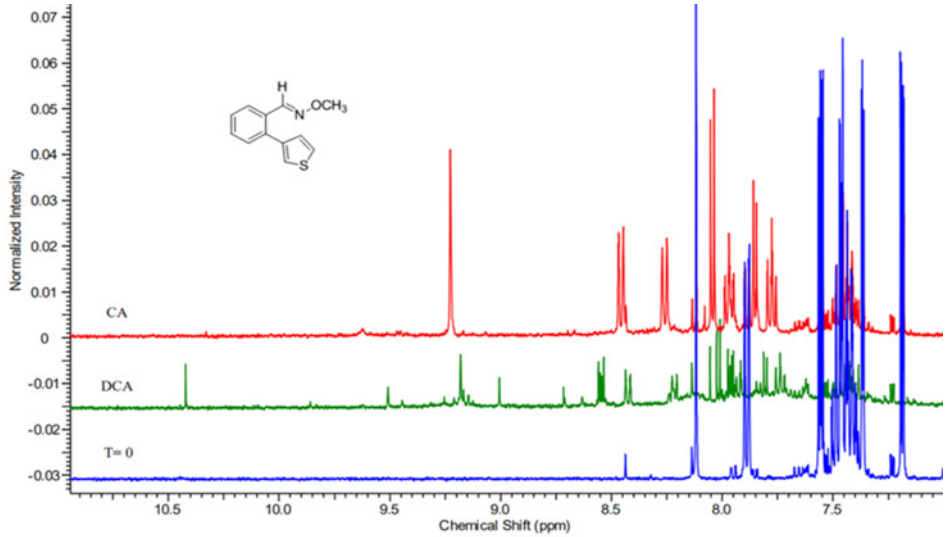


Figure 5. *Overlaid ^{1}H NMR spectra displaying the photolysis reaction of (E)-2-(thiophen-3-yl)benzaldehyde O-methyl oxime at t=0 and t=8 hours for both Chloranil (CA) and 9,10-dicyanoanthracene (DCA).*

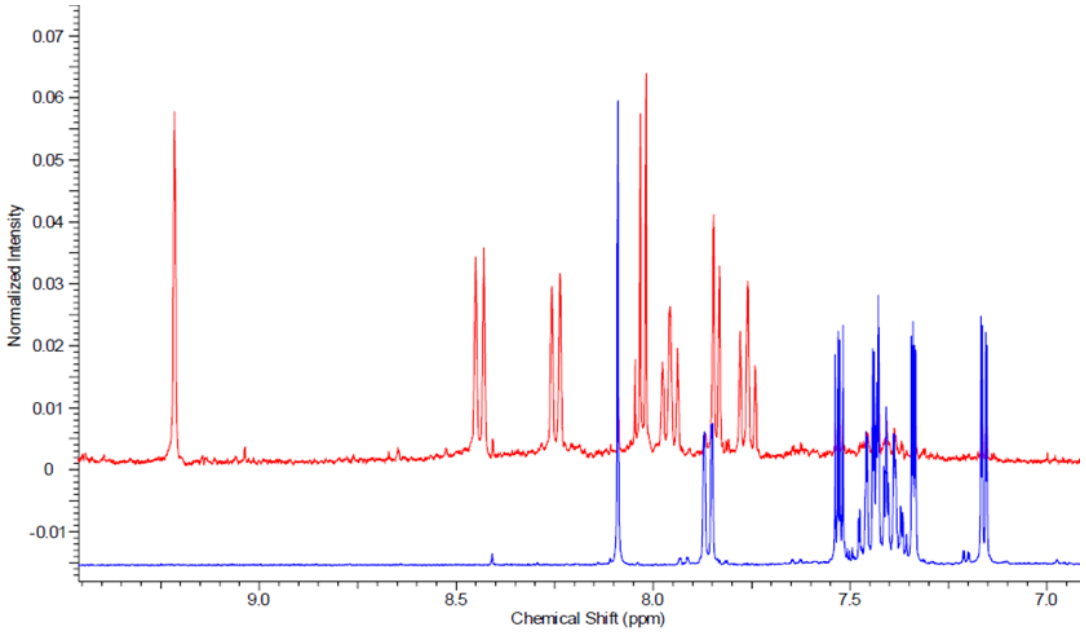


Figure 6. *Overlaid ^{1}H NMR spectra displaying the photolysis reaction of (E)-2-(thiophen-3-yl)benzaldehyde O-methyl oxime at t=0 and t=8 hours for chloranil (CA).*

The product from the reaction of (*E*)-2-(thiophen-3-yl)benzaldehyde O-methyl oxime in the presence of CA was positively identified as thieno[2,3-c]isoquinoline when it was isolated and analyzed by NMR (Figure 7). The other possible cyclic product, thieno[3,4-c]isoquinoline, was not observed, suggesting that this reaction is regioselective and therefore a potentially useful method for preparing such products selectively.

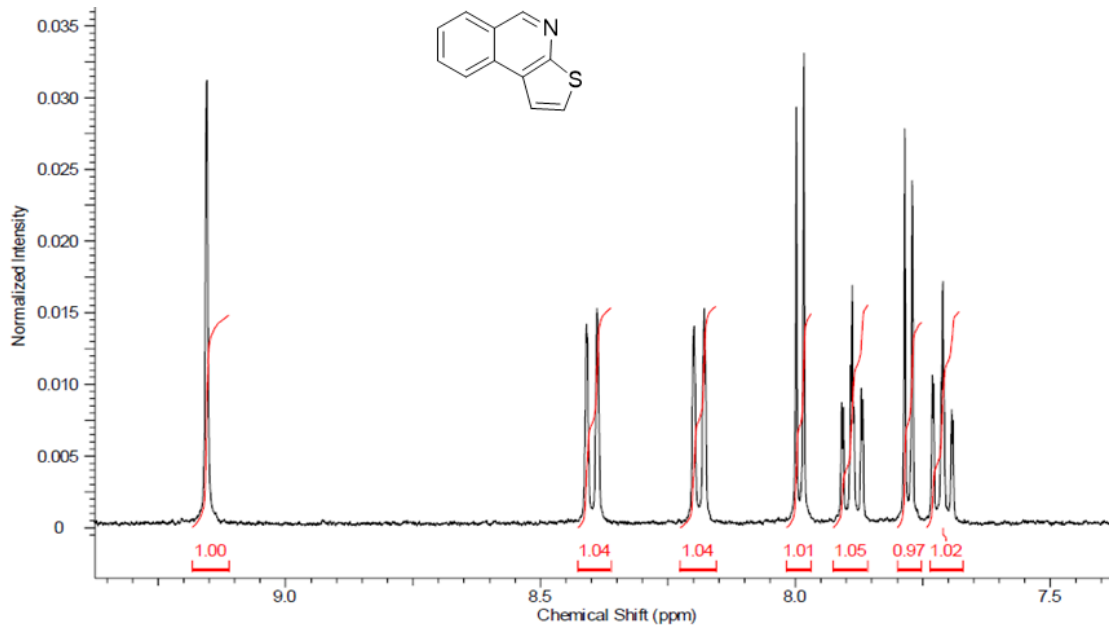


Figure 7. *Isolated product from the reaction of (E)-2-(thiophen-3-yl)benzaldehyde O-methyl oxime with Chloranil (CA).*

Furan Derivatives

Irradiation of the two furan derivatives, (*E*)-2-(furan-2-yl)benzaldehyde O-methyl oxime and (*E*)-2-(furan-3-yl)benzaldehyde O-methyl oxime, under the same conditions as those of the thiophene derivatives, did not yield any cyclic products as the major pathway. Photolysis of (*E*)-2-(furan-2-yl)benzaldehyde O-methyl oxime with CA seemed to result in *E/Z* isomeration (Figure 8; red line). After photolysis, the imine proton remains visible but other peaks in the aromatic region had shifted slightly, which is typical for *E/Z* isomers. This result is similar to that observed for the photolysis of the thiophene derivative under the same conditions. Irradiation of (*E*)-2-(furan-2-yl)benzaldehyde O-methyl oxime in the presence of DCA, however, yielded a mixture of products, none of which could be positively identified. Methanol formation was observed in the aliphatic region of the NMR spectrum, suggesting that cyclization is one of the pathways present under these conditions.

Unlike the reactions of (*E*)-2-(thiophen-3-yl) benzaldehyde O-methyl oxime, photolysis of (*E*)-2-(furan-3-yl)benzaldehyde O-methyl oxime in the presence of either CA or DCA gave little, if any, cyclization (Figure 9). In both cases, a variety of products are formed upon irradiation, however, it is unclear whether the desired cyclization is one of the pathways. Some methanol formation was observed, suggesting the cyclization pathway is present, however, the amounts are very small.

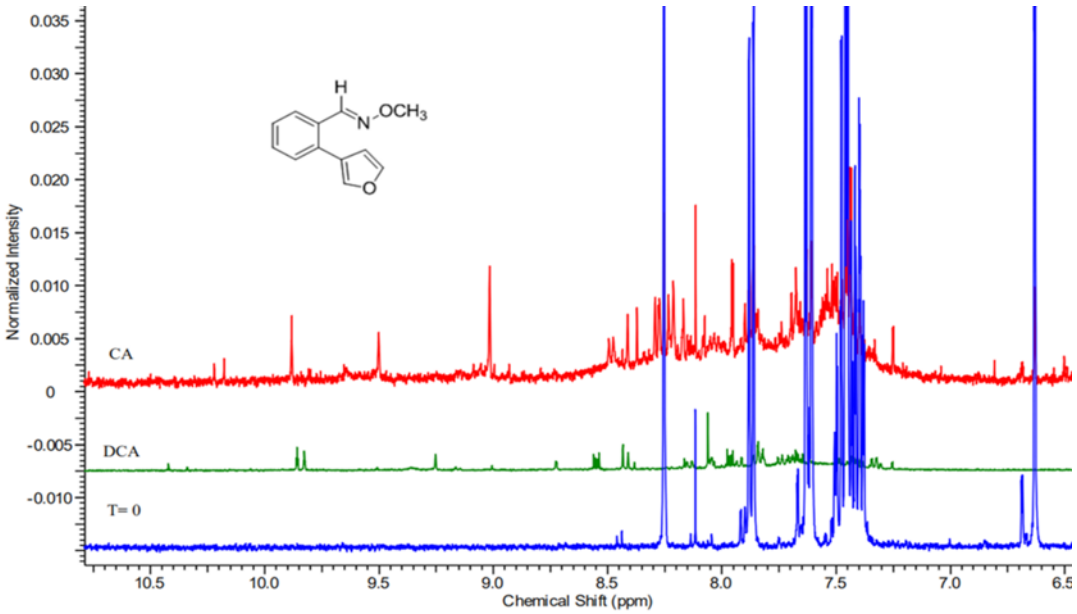


Figure 8. *Overlaid ^{1}H NMR spectra displaying the photolysis reaction of (E)-2-(furan-2-yl)benzaldehyde O-methyl oxime at t=0 and t=8 hours for both Chloranil (CA) and 9,10-dicyanoanthracene (DCA).*

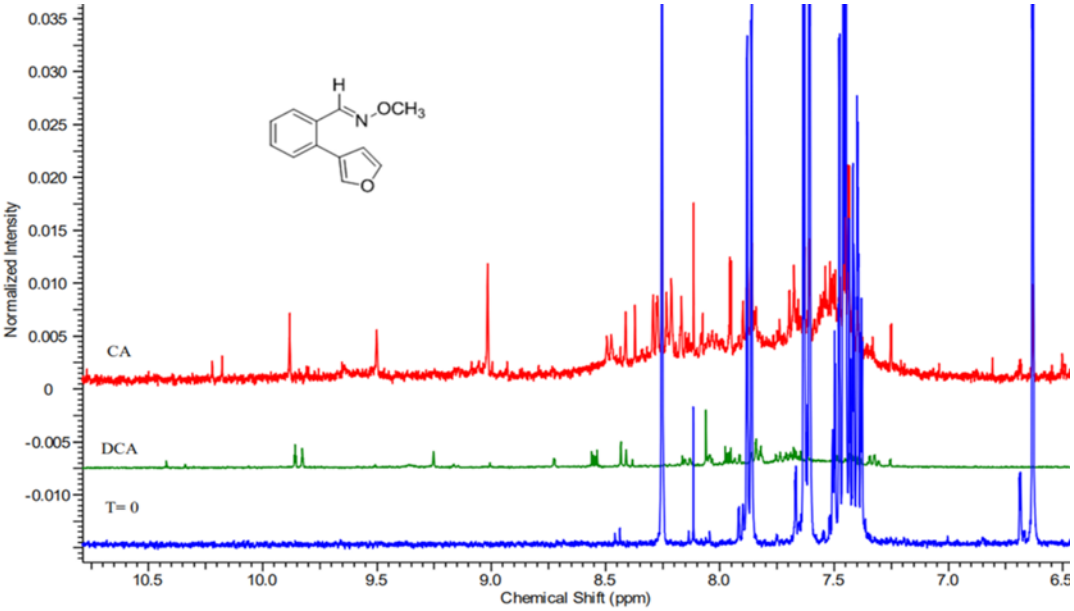


Figure 9. *Overlaid ^1H NMR spectra displaying the photolysis reaction of (E)-2-(furan-3-yl)benzaldehyde O-methyl oxime at t=0 and t=8 hours for both Chloranil (CA) and 9,10-dicyanoanthracene (DCA).*

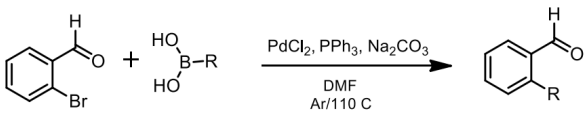
CONCLUSIONS

This study has shown that the use of certain heteroaromatic nucleophiles in photooxidative cyclization reactions, can lead to the formation of interesting heteroaromatic structures. Reacting (*E*)-2-(thiophen-2-yl)benzaldehyde O-methyl oxime with DCA or (*E*)-2-(thiophen-3-yl)benzaldehyde O-methyl oxime with CA resulted in the formation of thieno[3,2-c] isoquinoline and as thieno[2,3-c]isoquinoline respectively in very good yields. The reaction of (*E*)-2-(thiophen-3-yl)benzaldehyde O-methyl oxime with CA was also found to be regioselective; only one of two possible products was formed. Similar reactions with furans as nucleophiles did not give the same results. The reasons for these differences are unclear, although it can be hypothesized that the presence of sulfur increases the nucleophilicity of the ring structure. Further work is underway, including theoretical calculations on relative energies of the proposed intermediates to further learn about the mechanistic aspects of these interesting reactions.

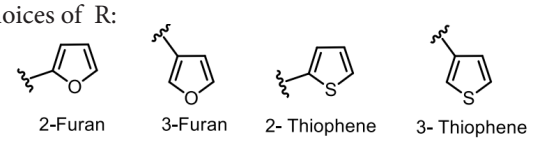
EXPERIMENTAL SECTION

Synthesis of 2-Aryl Heteroaromatic Ring Benzaldehydes

The 2-aryl heteroaromatic rings benzaldehyde derivatives were prepared by means of a Suzuki–Miyaura coupling reaction in which 2-bromobenzaldehyde is coupled with the boronic acid derivatives of the chosen heteroaromatic ring substrates as shown in Scheme 6. A pressure tube reactor (35 mL) supplied with a magnetic stir bar and purged by argon gas, was charged with triphenylphosphine (0.15 mol equiv), palladium (II) chloride (0.05 mol equiv), 2-bromobenzaldehyde (1 mol equiv), a boronic acid derivative of a heteroaromatic ring substrate (1 mol equiv), and 20 mL dimethylformamide (DMF).



Choices of R:



Scheme 6. *A diagram shows the Suzuki-Miyaura coupling reaction for 2-bromobenzaldehyde with the boronic acid derivative of heteroaromatic ring substrate (R).*

The reaction mixture was stirred at 110˚C for 10 minutes after which 4 mL of aqueous sodium carbonate (2M) was added to the mixture. The reaction was left running for 48 h under argon atmosphere. Thin-layer chromatography (TLC) was used to monitor the reaction. Subsequently, the mixture was extracted with diethyl ether (40 mL x 3), water (40 mL x 3), brine solution (40 mL x 3), and aqueous lithium chloride (5%) (40 mL x 3). The organic layer was dried over anhydrous magnesium sulfate, and then concentrated in vacuo. The concentrated mixture was purified by centrifugal chromatography using a chromatotron with a gradient of ethyl acetate and hexane as the eluting solvent. The synthesized parent aldehydes, their isolated yields, and their 1HNMR data are reported below.

2-(furan-2-yl) benzaldehyde (1)

Using the general procedure described above, 2-bromobenzaldehyde (0.466 mL), 2-furan boronic acid (0.4899g), palladium (II) chloride (0.0417g), and triphenylphosphine (0.1972g) were reacted to yield the desired aldehyde. Purification was achieved by using a 10% ethyl acetate - 90% hexane solvent mixture. Yield: 0.7173g (98%), yellow oil; ^{1}H-NMR (400 MHz, ACETONITRILE-d3) δ ppm 6.63 (d, J=5.26 Hz, 1 H) 6.78 (d, J=3.42 Hz, 1 H) 7.46 - 7.52 (m, 1 H) 7.66 - 7.76 (m, 4 H) 7.91 (d, J=7.82 Hz, 1 H) 10.33 (s, 1 H).[16]

2-(furan-3-yl) benzaldehyde (2)

Using the general procedure described above, 2-bromobenzaldehyde (0.466 mL), 3-furan boronic acid (0.4866g), palladium (II) chloride (0.04200g), and triphenylphosphine (0.1951g) were reacted to yield the desired aldehyde. Purification was achieved by using a 10% ethyl acetate, 90% hexane solvent mixture. Yield: 0.6605g (95%), Dark yellow oil: ^{1}H-NMR (400 MHz, ACETONITRILE-d3) δ ppm 6.71 (s, 1 H) 7.49 - 7.56 (m, 2 H) 7.62 - 7.71 (m, 3 H) 7.93 (d, J=7.83 Hz, 1 H) 10.19 (s, 1 H).[17]

2-(thiophen-2-yl) benzaldehyde (3)

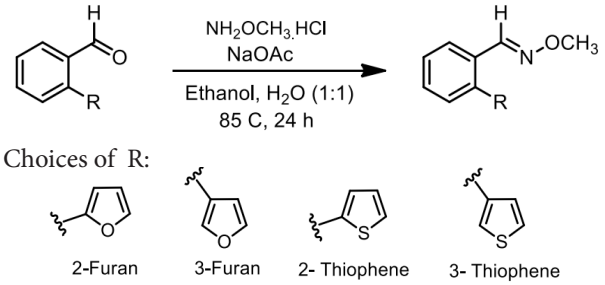
Using the general procedure described above, 2-bromobenzaldehyde (0.466 mL), 2-thiophene boronic acid (0.5648g), palladium (II) chloride (0.0506g), and triphenylphosphine (0.2028g) were reacted to yield the desired aldehyde. Purification was achieved by using a 20% ethyl acetate, 80% hexane solvent mixture. Yield: 0.8338g (98%), brown oil. ^{1}H-NMR (400 MHz, ACETONITRILE-d3) δ ppm 7.15 - 7.22 (m, 2 H) 7.51 - 7.64 (m, 4 H) 7.66 - 7.72 (m, 1 H) 7.95 (d, J=7.82 Hz, 1 H) 10.13 (s, 1 H).[16]

2-(thiophen-3-yl) benzaldehyde (4)

Using the general procedure described above, 2-bromobenzaldehyde (0.466 mL), 3-thiophene boronic acid (0.51694g), palladium (II) chloride (0.04680g), and triphenylphosphine (0.18061g) were reacted to yield the desired aldehyde. Purification was achieved by using a 20% ethyl acetate, 80% hexane solvent mixture. Yield: 0.7311g (97%), brown oil. ^{1}HNMR (400 MHz, ACETONITRILE-d3) δ ppm 7.27 (d, J=5.01 Hz, 1 H) 7.46 (d, J=4.28 Hz, 1 H) 7.50 - 7.59 (m, 3 H) 7.64 - 7.72 (m, 1 H) 7.94 (d, J=7.82 Hz, 1 H) 10.07 (s, 1 H).[18]

Synthesis of 2-Aryl Heteroaromatic Ring Benzaldehyde Oxime Ethers

In a 100 mL pressure tube reactor equipped with a magnetic stir bar, a solution of the appropriate 2-aryl heteroaromatic ring benzaldehyde **(1–4)** (1 mol equiv) was dissolved in a 1:1 mixture of ethanol and water (60 mL total volume). Next, methoxylamine hydrochloride (2 mol equiv) and sodium acetate (4.5 mol equiv) were added to the mixture and the reaction was stirred for 24 hours at 85˚C. This is shown in Scheme 7 below.



Scheme 7. *A diagram shows the synthetic procedure of making 2-aryl heteroaromatic ring benzaldehyde oxime ethers.*

The mixture was then extracted with diethyl ether (40 mL x 3), water (40 mL x 3), and brine solution (40 mL x 3). The organic layer was dried over anhydrous magnesium sulfate, and then concentrated in vacuo. Centrifugal chromatography (Chromatotron) was used to purify the oxime ether using a gradient of ethyl acetate and hexane.

(*E*)-2-(furan-2-yl) benzaldehyde O-methyl oxime (5)

Using the general procedure described above, 2-(furan-2-yl) benzaldehyde (0.6887g), methoxyamine hydrochloride (0.6681g), sodium acetate (1.476g), and ethanol with water (1:1) 60 mL were reacted to prepare the desired oxime ether. Purification of the crude material was achieved using a 10% ethyl acetate - 90% hexane solvent mixture. Yield: 0.6817g (85%), brown oil. ^{1}H-NMR: (400 MHz, ACETONITRILE-d3) δ ppm 3.93 (s, 3 H) 6.58 (s, 2 H) 7.33 - 7.40 (m, 1 H) 7.44 - 7.50 (m, 1 H) 7.64 (t, J=3.91 Hz, 2H) 7.83 (dd, J=7.89, 1.41 Hz, 1 H) 8.42 (s, 1 H). 13C-NMR: (101 MHz, CDCl3) δ 62.01, 110.01, 111.57, 127.07, 127.89, 127.92, 128.85, 129.63, 130.30, 142.98, 148.26, 152.07.

(*E*)-2-(furan-3-yl) benzaldehyde O-methyl oxime (6)

Using the general procedure described above, 2-(furan-3-yl) benzaldehyde (1.0823g), methoxyamine hydrochloride (1.06020g), sodium acetate (2.349g), and ethanol with water (1:1) 60 mL were reacted to prepare the desired oxime ether. Purification of the crude material was achieved using a 10% ethyl acetate, 90% hexane solvent mixture. Yield: 1.0766g (85%), light brown oil. ^{1}H-NMR: (400 MHz, ACETONITRILE-d3) δ ppm 3.91 (s, 3 H) 6.60 (s, 1 H) 7.33 - 7.46 (m, 3 H) 7.59 (dd, J=9.66, 2.93 Hz, 2 H) 7.84 (d, J=7.82 Hz, 1 H) 8.22 (s, 1 H). 13C-NMR: (101 MHz, CDCl3) δ 61.95, 111.86, 123.89, 126.42, 127.57, 129.67, 129.81, 130.02, 132.63, 140.59, 143.14, 147.82.

(*E*)-2-(thiophen-2-yl) benzaldehyde O-methyl oxime (7)

Using the general procedure described above, 2-(thiophen-2-yl) benzaldehyde (0.6916g), methoxyamine hydrochloride (0.6764g), sodium acetate (1.4073g), and ethanol with water (1:1) 60 mL were reacted to prepare the desired oxime ether. Purification of the crude material was achieved using a 20% ethyl acetate, 80% hexane solvent mixture. Yield: 1.0766g (98%), yellow oil. ^{1}H-NMR: (400 MHz, ACETONITRILE-d3) δ ppm 3.92 (s, 3 H) 7.06 (d, J=3.55 Hz, 1 H) 7.15 (dd, J=5.14, 3.55 Hz, 1 H) 7.39 - 7.50 (m, 3 H) 7.51 - 7.54 (m, 1 H) 7.87 (dd, J=7.64, 1.53 Hz, 1 H) 8.20 (s, 1 H). 13C-NMR: (101 MHz, CDCl3) δ 61.99, 126.32, 126.57, 127.46, 128.05, 128.08, 129.53, 130.32, 130.79, 134.34, 140.67, 147.81.

(*E*)-2-(thiophen-3-yl) benzaldehyde O-methyl oxime (8)

Using the general procedure described above, 2-(thiophen-3-yl) benzaldehyde (0.6809g), methoxyamine hydrochloride (0.6497g), sodium acetate (1.356g), and ethanol with water (1:1) 60 mL were reacted to prepare the desired oxime ether. Purification of the crude material was achieved using a 20% ethyl acetate, 80% hexane solvent mixture. Yield: 0.7711g (98%), dark yellow oil. ^{1}H-NMR: (400 MHz, ACETONITRILE-d3) δ ppm 3.90 (s, 3 H) 7.15 - 7.17 (m, 1 H) 7.34 (d, J=4.40 Hz, 1 H) 7.36 - 7.48 (m, 3 H) 7.51 - 7.55 (m, 1 H) 7.86 (d, J=8.56 Hz, 1 H) 8.09 (s, 1 H). 13C-NMR: (101 MHz, CDCl3) δ 61.94, 123.93, 125.77, 126.25, 127.59, 129.10, 129.57, 129.94, 130.00, 136.70, 139.97, 147.91.

Photolysis of Oxime Ethers.

A solution (10-15 mM) of the oxime ether and acetonitrile-d3 (CD3CN) was made and 1 mL of the solution was added to each nuclear magnetic resonance (NMR) tubes. A photosensitizer (9,10-dicyanoanthracene or chloranil) (5 mmol) was added to the NMR tubes. The samples prepared with 9,10-dicyanoanthracene (DCA) were photolyzed at 420 nm using a photoreactor for up to 8 hours. The samples prepared with chloranil (CA) were photolyzed at 350 nm using a photoreactor for up to 8 hours. The reaction was monitored by ^{1}H-NMR every hour to track product production. An internal standard of either 4-nitrobenzyl-bromide or 4-nitrobenzaldehyde was added following photolysis to determine product yield by 1H-NMR. Gas chromatography-mass spectrometry (GC-MS) was also used to verify product formation.

REFERENCES

1. Ritz, J.; Fuchs, H. *Ullmann's Encyclopedia of Industrial Chemistry*, Moran (2005).
2. Guengerich, F. P. *Chem. Res. Toxicol.*, **2001**, *14*, 611.
3. (a) Beckmann, E. *Ber. Bunsen-Ges. Phys. Chem.,* **1886**, *19*, 988. (b) Gawley, R. E. *Org. Reactions* **1988**, *35*, 14–24. (c) Gregory, B. J.; Moodie, R. B.; Schofield, K. *J. Chem. Soc. (B)*, **1970**, 338.
4. (a) Bieleman, J. H.; Bolle, T.; Braig, A.; Glaser, J. K.; Spang, R.; Köhler, M.; Valet, A. *in Additives for Coatings* Bieleman, J. H. Ed. Wiley-VCH, Weinheim, **2000**, 257-348. (b) van Gorkum, R.; Bouwman, E. *Coord. Chem. Rev.*, **2005**, *249*, 1709.
5. Tanase, S.; Hierso, J.-C.; Bouwman, E.; Reedijk, J.; ter Borg, J.; Bieleman, J. H.; Schut, A. *New J. Chem.* **2003**, *27*, 854.
6. Guengerich, F.P.; Macdonald, T.L. *FASEB J.,* **1990**, *4*, 2453.
7. (a) de Lijser, H. J. P.; Tsai, C. K. *J. Org. Chem.* **2004**, *69*, 3057. (b) de Lijser, H. J. P.; Rangel, N. A.; Tetalman, M. A.; Tsai, C. K. *J. Org. Chem.* **2007**, *72*, 4126.
8. Hofstra, J. L.; Grassbaugh, B. R.; Tran, Q. M.; Armada, N. R.; de Lijser, H. J. P. *J. Org. Chem.* **2015**, 80 (1), 256–265.
9. Kim, W. S. Photoinduced radical cyclization of O-arylalkynyl oximes; CSUF. 2013, Master Thesis.
10. Miyaura, N.; Yamada, K.; Suzuki, A. *Tetrahedron Lett.* **1979**, *20,* 3437.
11. Ye, F.; Shi, Y.; Zhou, L.; Xiao, Q.; Zhang, Y.; Wang, J. *Org. Lett.* **2011**, *13*, 5020.
12. de Lijser, H. J. P.; Fardoun, F. H.; Sawyer, J. R.; Quant, M. *Org. Lett.* **2002**, *4*, 2325.
13. (a) de Lijser, H. J. P.; Kim, J. S.; McGrorty, S. M.; Ulloa, E. M. *Can. J. Chem.* **2003**, *81*, 575. (b) Park, A.; Kosareff, N. M.; Kim, J. S.; de Lijser, H. J. P. *Photochem. Photobiol*. **2006**, *82*, 110.
14. de Lijser, H. J. P.; Hsu, S.; Marquez, B. V.; Park, A.; Sanguantrakun, N.; Sawyer, J. R. *J. Org. Chem.* **2006**, *71*, 7785.
15. de Lijser, H. J. P.; Burke, C. R.; Rosenberg, J.; Hunter J. *J. Org. Chem.* **2009**, *74*, 1679.
16. Jordan, A. D.; Orsini, M. J.; Middleton, S. A.; Connolly, P. J.; Brenneman, D. E.; Pan, K.; Reitz, A. B. 8-(heteroaryl) phenalkyl-1-phenyl-1,3,8-triazaspiro[4.5]decan-4-ones as opioid receptor modulators. Med. Chem. 2005, 1, 601-610. (39)
17. Zhao, J.; Yue, D.; Campo, M. A.; Larock, R. C. An Aryl to Imidoyl Palladium Migration Process Involving Intramolecular C-H Activation. J. Am. Chem. Soc. 2007, 129, 5288-5295
18. Hartman, G. D.; Halczenko, W.; Phillips, B. T. Iminium ion mediated cyclizations with 4-aryl-1,4-dihydropyridines. Bridging with thiophene and furan. J. Org. Chem. 1986, 51, 142-148.

Development and Evaluation of a Homology Model of Human Fatty Acid Amide Hydrolase Enzyme

Mark Rodriguez
Advisor: Stevan Pecic, Ph.D.
Department of Chemistry and Biochemistry, California State University, Fullerton

ABSTRACT

The purpose of this study is to create a valid homology model for the human fatty acid amide hydrolase (FAAH) enzyme using as a 3D template a known experimental structure of the rat fatty acid amide hydrolase enzyme. This study was done by utilizing resources such as the protein database, ICM Pro molecular modeling software, and a structural analysis verification server. From observing the results of the five criteria from the verification server, it can be suggested that a valid homology model for human fatty acid amide hydrolase was created. Structural information provided by this homology model will be valuable in structure-based drug discovery and drug design of future FAAH inhibitors.

INTRODUCTION

Fatty acid amide hydrolase (FAAH) is an integral membrane enzyme that hydrolyzes the endocannabinoids, which are endogenous lipid ligands that activate the cannabinoid receptors, CB1 and CB2.[1] Two endocannabinoids have been identified in mammals, anandamide (AEA) and 2-arachidonoylglycerol (2-AG). FAAH terminates AEA signaling by hydrolyzing AEA to arachidonic acid and ethanolamine; two metabolites that do not activate cannabinoid receptors (Figure 1).[2]

The FAAH enzyme possesses an unusual serine-serine-lysine (Ser241-Ser217-Lys142) catalytic triad that is conserved among different species. The first amino acid in this catalytic triad, Ser241 acts as a nucleophile and Lys142 is serving as an acid/base catalyst during hydrolysis of the amide bond. Ser217 serves as a 'proton shuttle' between Ser241 and Lys142, allowing FAAH to exhibit an unusual ability to hydrolase amides and esters at equivalent rates.

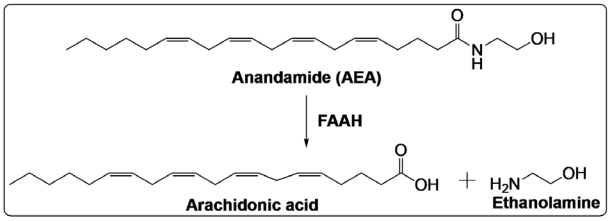


Figure 1. *Anandamide is hydrolyzed to arachidonic acid and ethanolamine by FAAH*

The pharmacological inactivation of FAAH produces analgesic, anti-inflammatory, anti-fibrotic, anxiolytic and mild antidepressant effects, without showing the undesirable side effects of direct cannabinoid receptor agonists. This indicates that FAAH is a promising therapeutic target for pain management and various pulmonary diseases.[1] Several classes of FAAH inhibitors have been reported, including carbamates, ureas, and substrate-derived inhibitors (Figure 2).[2, 3] However, the most potent among these inhibitors are mostly irreversible inhibitors that non-selectively inhibit other hydrolases in peripheral tissues, including several carboxylesterases.[4, 5] There is a clearly an unmet need to find novel, potent FAAH inhibitors that will be used as a potential therapeutics in pain management.

As part of our ongoing efforts to develop novel FAAH inhibitors as potential therapeutics in pain management, we plan to design and synthesize a series of analogs based on the structure of known FAAH inhibitor PF750 (Figure 2) and evaluate their inhibition

for human FAAH enzyme. In addition, in order to better understand the binding modes of novel FAAH inhibitors and to complement future drug design efforts, we plan to conduct a docking study with our library of compounds. Since the crystal structure of the human FAAH enzyme has not yet been solved, our approach necessitates the use of a homology modeling paradigm. The overall goal of the experiment was to create an optimal and valid homology model for human fatty acid amide hydrolase enzyme and to identify the binding pocket in order to utilize it for docking studies in future experiments.

Here we describe our efforts to prepare a homology model of human FAAH enzyme and its validation.

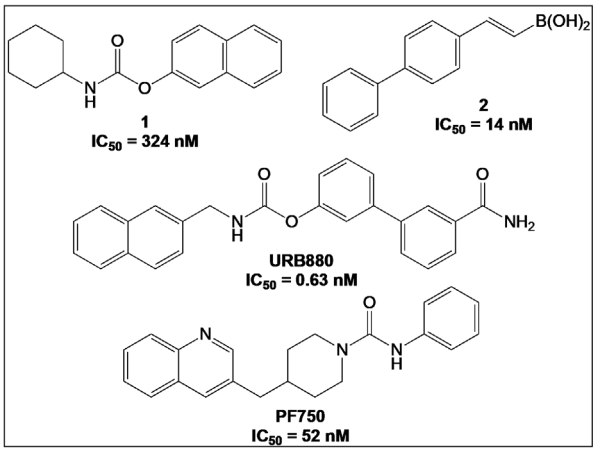


Figure 2. *Known FAAH Inhibitors*

METHODS

Sequence Alignment, Homology Model Building and Model Evaluation

Amino acid sequence of human FAAH enzyme was retrieved from NCBI protein database.[6] Sequence alignment was carried out with the ICM Pro 3.61 Pro (based on ZEGA sequence alignment[7] - Needleman and Wunsch algorithm with zero gap and penalties). To make the homology model, we used a homology algorithm from ICM Pro 3.6. Making a homology modeling ICM PRO required an initial placement of the aligned polypeptide chain to the template structure. During this process the side-chain torsion angles are predicted through optimization simulations for the energy of all non-identical residues. The conformation model of protein secondary structure relies heavily on the internal coordinate definition and Biased Probability Monte Carlo (BPMC) optimization. The analysis of how well the 3D model appears is calculated by ICM Pro, along with creating possible backbone deviations between homologues in the model. Between several crystal structures available in PDB database, we selected as a template a crystal structure of rat FAAH enzyme (PDB code: 3QK5). In order to create the suitable homology model for human FAAH, the sequence of human FAAH and the rat FAAH PDB template were converted to the ICM model. Conversions to ICM model were performed by optimization of hydrogens, certain amino acids (H, P, N, C, Q), and ligand tautomer, assignment of the secondary structure, and approximations of the formal charges of the ligands. This allowed for the new structure to be formed using the rat FAAH enzyme crystal structure as the template. Following the completion of the building the homology model, we identified the binding pocket. In order to validate the model, programs such as Procheck, Verify3D, Errat, What check and Prove from SAVES metaserver were used.

RESULTS AND DISCUSSION

Homology modeling has become a common technique in the last couple years, and this procedure usually consists of four sequential steps:[8]
1) template recognition; 2) alignment of the target sequence and template structure(s); 3) model building; and 4) model verification.
Docking studies usually follow homology modeling as separate experiments.

1) Identification of suitable template- The main challenge in homology modeling is to find a suitable template. Since the crystal structure of the human FAAH enzyme is still not solved, the only way to achieve our objective is to use the most similar enzyme structure (that has been determined by X-ray crystallography) as a template

to build the homology model. The first step of any homology modeling method begins with the selection of a suitable structural template from the Protein Data Base (http://www.pdb.org).[9] As a template for building the human FAAH enzyme homology model, we selected the crystal structure of the rat fatty acid amide hydrolase (PDB code: 3QK5) since its crystal structure is solved at the highest resolution (2.20 Å) that we found deposited. Generally, values of crystal structures of 2.2 Å and below are optimal for homology modeling and docking studies.

2) Sequence alignments- Despite recent advances in methods used to generate homology models, a difficult aspect is the initial alignment between the target sequence and template structure. A common method to align the target and the template is through usage of a sequence alignment software. Alignments of target sequence to the template structures were made using the ICM Pro sequence alignment program, using program default parameters. The ICM Pro alignment program is based on a ZEGA algorithm (zero gap and penalties). The sequence alignment of the human FAAH enzyme shows excellent sequence identity with the rat FAAH enzyme sequence template with more than 80% sequence similarity (Figure 3). As seen in Figure 3, the alignment of rat and human FAAH sequences contains many conserved residues (shown by green) that are essential for enzyme function. After obtaining the sequence alignment, a homology model can be built using rat FAAH PDB file as the template.

3) Homology modeling of a human FAAH enzyme- After obtaining a suitable template and generating an alignment between the target (human FAAH enzyme) and the template structure (rat FAAH enzyme), spatial restraints were used to create a homology model of the human FAAH enzyme (Figure 4). Spatial restraints can be obtained from a number of sources that include homology-derived restraints and molecular mechanics-based or statistically derived preferences for bonded and non-bonded interactions.[8] An example of a program that uses this style of homology modeling is ICM Pro. In addition to the satisfaction of spatial restrains, another successful method to generate homology models is the concept of positional tethers. Tethering and other proprietary methods are also available in ICM Pro. From doing this, a model can be formed. Figure 4 depicts the outcome of the modeling; where each component of the secondary structure in the model was colored in order to better differentiate each shape. The homology model was then reconverted back to the PCB form in order to assess validity of the model.

```
80% pP=69.9                                 ....#.#.##.##....G..#...##......RW......GR..ARGA#.RARQ+QRA.LE.MD+A#QRFRLQNPDLDSEALL.LPL#QLVQKL
3QK5_A_PDBID_CHAIN_SEQUENCE     1   MGGSHHHHHHGMASMTGGQQMGRDLYDDDDKDRWGSELETGRQKARGAATRARQKQRASLETMDKAVQRFRLQNPDLDSEALLTLPLLQLVQKL
NP_001432_2                     1   ---MVQYELWAALPGASGVALACCFVAAAVALRW-----SGRRTARGAVVRARQRQRAGLENMDRAAQRFRLQNPDLDSEALLALPLPQLVQKL

                                    .S.EL.PEAV#FTY#GKAWEVNKGTNCVTSYL.DCETQLSQAPRQGLLYGVPVSLKECF.YKG.DSTLGLSLNEG#P.E.D.VVV.VLKLQGAV
3QK5_A_PDBID_CHAIN_SEQUENCE    95   QSGELSPEAVFFTYLGKAWEVNKGTNCVTSYLTDCETQLSQAPRQGLLYGVPVSLKECFSYKGHDSTLGLSLNEGMPSESDCVVVQVLKLQGAV
NP_001432_2                    87   HSRELAPEAVLFTYVGKAWEVNKGTNCVTSYLADCETQLSQAPRQGLLYGVPVSLKECFTYKGQDSTLGLSLNEGVPAECDSVVVHVLKLQGAV

                                    PFVHTNVPQSM#S#DCSNPLFGQT#NPWKSSKSPGGSSGGEGALIGSGGSPLGLGTDIGGSIRFPS.FCGICGLKPTGNRLSKSGLKGCVYGQ.
3QK5_A_PDBID_CHAIN_SEQUENCE    189  PFVHTNVPQSMLSFDCSNPLFGQTMNPWKSSKSPGGSSGGEGALIGSGGSPLGLGTDIGGSIRFPSAFCGICGLKPTGNRLSKSGLKGCVYGQT
NP_001432_2                    181  PFVHTNVPQSMFSYDCSNPLFGQTVNPWKSSKSPGGSSGGEGALIGSGGSPLGLGTDIGGSIRFPSSFCGICGLKPTGNRLSKSGLKGCVYGQE

                                    AV.LS#GPMARDVESLALCL+ALLCE.#F.LDPTVPPLPFREEVY.SS.PLRVGYYETDNYTMPSPAMRRA##ETKQ.LEAAGHTL#PFLP.NI
3QK5_A_PDBID_CHAIN_SEQUENCE    283  AVQLSLGPMARDVESLALCLKALLCEHLFTLDPTVPPLPFREEVYRSSRPLRVGYYETDNYTMPSPAMRRALIETKQRLEAAGHTLIPFLPNNI
NP_001432_2                    275  AVRLSVGPMARDVESLALCLRALLCEDMFTLDPTVPPLPFREEVYTSSQPLRVGYYETDNYTMPSPAMRRAVLETKQSLEAAGHTLVPFLPSNI

                                    P#ALE.LS.GGLFSDGG..FLQNFKGDFVDPCLGDL#.IL+LP.W#K.LL.#L#KPL#PRL.AFL..M+.RSA.KLW.LQHEIE#YR..VIAQW
3QK5_A_PDBID_CHAIN_SEQUENCE    377  PYALEVLSAGGLFSDGGRSFLQNFKGDFVDPCLGDLILILRLPSWFKRLLSLLLKPLFPRLAAFLNSMRPRSAEKLWKLQHEIEMYRQSVIAQW
NP_001432_2                    369  PHALETLSTGGLFSDGGHTFLQNFKGDFVDPCLGDLVSILKLPQWLKGLLAFLVKPLLPRLSAFLSNMKSRSAGKLWELQHEIEVYRKTVIAQW

                                    +A#.LDV#LTPML.PALDLN.PGRATGA#SYT#LYNCLDFPAGVVPVTTVTAED-AQME#Y+GYFGDIWD.#L.K.MK.SVGLPVAVQCVALPW
3QK5_A_PDBID_CHAIN_SEQUENCE    471  KAMNLDVLLTPMLGPALDLNTPGRATGAISYTVLYNCLDFPAGVVPVTTVTAEDDAQMELYKGYFGDIWDIILKKAMKNSVGLPVAVQCVALPW
NP_001432_2                    463  RALDLDVVLTPMLAPALDLNAPGRATGAVSYTMLYNCLDFPAGVVPVTTVTAEDEAQMEHYRGYFGDIWDKMLQKGMKKSVGLPVAVQCVALPW

                                    QEELCLRFMREVE.LMTP.KQ.S
3QK5_A_PDBID_CHAIN_SEQUENCE    565  QEELCLRFMREVEQLMTPQKQPS
NP_001432_2                    557  QEELCLRFMREVERLMTPEKQSS
```

Figure 3. *Sequence alignment of rat and human FAAH*

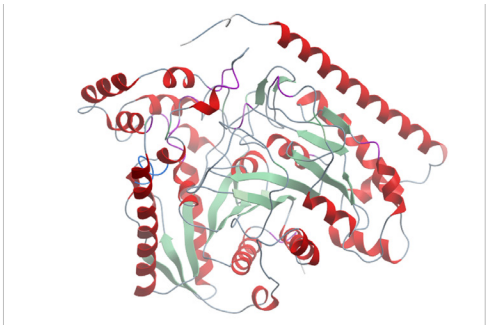


Figure 4. *Homology model of human FAAH*

4) Model evaluation- Evaluation methods verify whether a model satisfies standard steric and geometric criteria. All of the tools used in the construction of a model (e.g. template selection, alignment, model building, and refinement) has its own internal measures of validity. In order to evaluate the constructed human FAAH enzyme homology model for the purpose of our study, we used several different programs for evaluation that are available via the server of the UCLA-DOE Institute for Genomics and Proteomics. This Institute offers a service called Structural Analysis and Verification Server (SAVES)[10] that encompass five verification tools for the model evaluation: a) Procheck[11]; b) What_check[10]; c) Errat[12]; d) Verify_3D[13]; and e) Prove[14].

a) A homology model was first evaluated using the program Procheck. The aim of Procheck is to assess how normal, or unusual, the geometry of the residues in a given protein structure is, as compared with stereochemical parameters derived from well-refined, high resolution structures. Ramachandran plot, which is a plot of the φ-ψ torsion angles for all residues in the structure, is one of the important parameters in Procheck for assessment of structures. Ramachandran plots for the human FAAH homology model is shown in Figure 5. A simple measure of quality that can be used from the plot is the percentage of residues in core region, and allowed regions should be very high (>90% residues). Another important factor in structural assessment is Goodness factor or G-Factor which shows

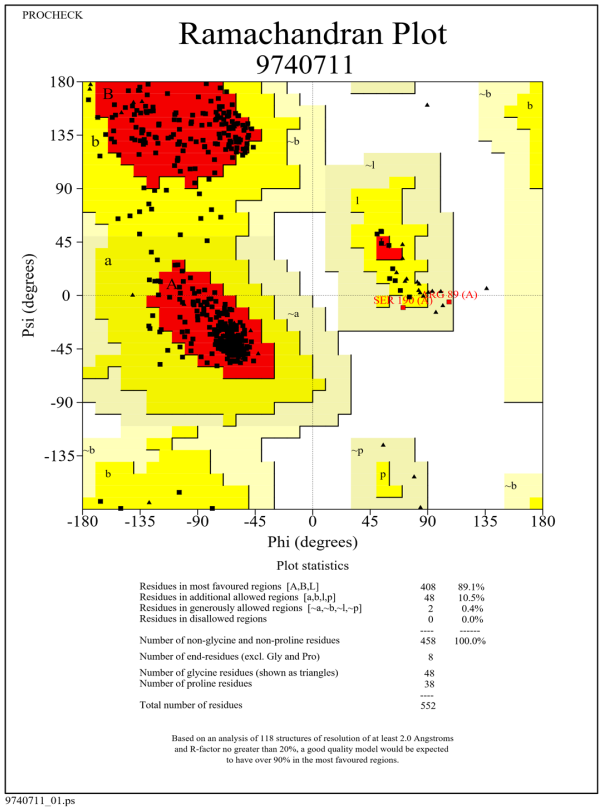


Figure 5. *Ramachandran plot of human FAAH homology model*

the quality of dihedral, covalent and overall bond angles. These scores should be above -0.5 for a reliable model. From observing the Ramachandran plot it can be seen that 89.1% of the residues are the most favorable region. In addition, 10.5%of the residues were found in the additionally allowed region. Although the percentage of residue in the most favorable region is not over 90%, it is close enough to deem it valid. To add on, the model was found to have a G-factor of 0.2, meaning the quality of the bond angles are good.

b) The results obtained from Errat are shown in Figure 6. This program counts the number of nonbonded interactions between atoms (CC, CN, CO, NN, NO, and OO) within a cutoff distance of 3.5 Å. It gives an overall quality factor for each protein structure, which is expressed as the percentage of protein for which the calculated error value falls below 95% rejection limit.

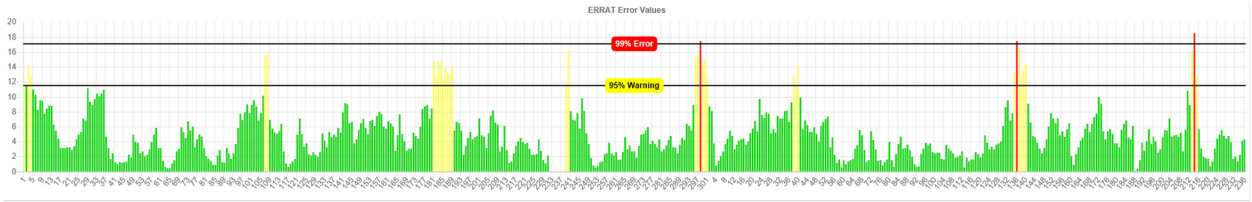


Figure 6. *Errat analysis of human FAAH homology model.*

Good structures produce these values above 90%. Accepted range is above 50%. The template structure of rat FAAH enzyme, PDB: 3QK5, has an Errat value of 85.294. From running an Errat on the FAAH homology model, it was found to have an overall quality factor of 93.585. Comparing the original template to the model shows that the model as a whole is more refined. In addition, a quality factor of 93.585 suggests that the structure of the model is valid.

c) The Prove results for the human FAAH homology model are shown in Figure 7. The stereochemical criterion implemented in this method is regularity or irregularity of atom volume. It provides an average volume Z-score of all the atoms. Z-score is calculated as the difference between the volume of the atom and the mean atomic volume for the corresponding atom type, divided by the standard deviation of the appropriate distribution.

Atoms in test proteins are scored using a Z-score standard deviation (Zstd), which is how many standard deviations their volume is from the mean (Zsm) for that atom type. A structural Z-score RMS (Zrms) is calculated using the square root of these scores. High scores have been found to be associated with uncertainty in the structure. Structures with poor resolution generally have a Zrms greater than 1.2, while for well resolved structures, the Zrms is around 1.0. The success of the Prove test was determine by the percentage of buried atoms in the model. In order to pass, less than 1% of the atoms should be buried. The caution range was between 1-5% while the error range is considered anything above 5%. From conducting the Prove test, it was found that the homology model had 5.1% of its atoms, as well as a total of 116 outlier protein atoms, buried. From this

Prove test, it can be determined that there were errors when creating the homology model. However, the model did show that the average Z score for the model for all resolutions hovered slightly above 0 but below 0.1 with the exception of one outlier at about 0.8. What this shows is that the certainty of the model is valid, but it contains an unusual number of buried atoms within it. The buried atoms are most likely hydrophobic residues within the protein. These buried atoms do not have direct contact with the ligand, but they are essential for

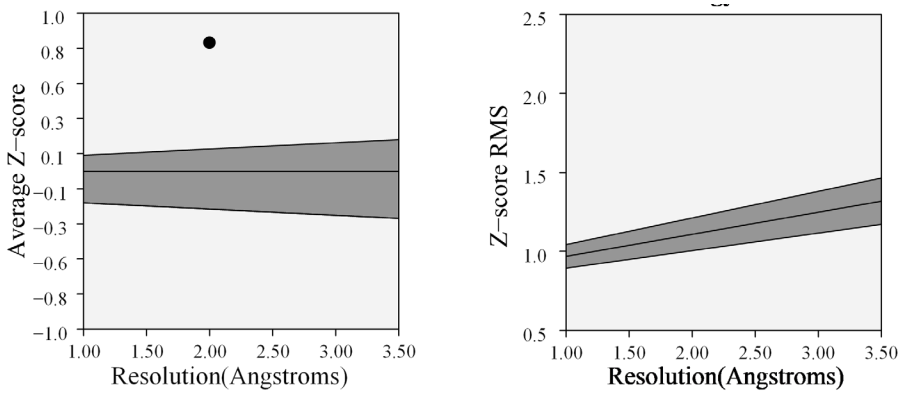


Figure 7. *Z-score results from Prove test of homology model.*

providing its shape and chemical properties. Changing the number of atoms buried within the protein can alter its three-dimensional shape to the point where it can possibly deactivate the active site. In order to keep the structural integrity of the homology model, the number of buried atoms were not altered.

d) Verify3D evaluates the sequences that are most compatible with the environments of the residues in the 3D structure. The environments are described by: the area of the residue buried in the protein and inaccessible to solvent, the fraction of sidechain area that is covered by polar atoms (O and N), and the local secondary structure. Verify3D works best on proteins with at least 100 residues. Verify3D measures a compatibility score to each residue of a full-atom protein structure, in other words, the compatibility of a model with its sequence. Verify3D results shown in Figure 8 represent the Verify3D average data score of the homology model generated in comparison with template. The score which matches the template and has value >0.2 for all or most of the residues indicates a reliable model. Negative or less than 0.2 scores are indicative of potential problems. From the data collected in the table, it was observed that 92.65% of the residues had averaged a 3D-1D score >= 0.2. An acceptable model should have 80% of its residues have a 3D-1D score >= 0.2. This data suggests that the 3D structure of the model is valid.

e) What check program is derived from a subset of protein verification tools from the WHATIF program. The program evaluates: bond lengths, bond angles, omega angle restraints, side chain planarity, improper dihedral distribution and inside/outside distribution. It does extensive checking of many stereochemical parameters of the residues in the model and it gives an overall summary of the quality of the structure as compared with current reliable structures. This summary is most useful for biologists seeking a good structure to use for modeling calculations. The first part of the Table 1 shows a number of constraint- independent quality indicators. All structure Z-scores, like, second generation packing quality, Ramachandran plot appearance, chi-1/ chi-2 rotamer normality, backbone conformation, and RMS Z-scores, should be positive. Positive values are just greater than average and do not necessarily mean that negative values are failure. The second part of these tables (in blue) mostly gives an impression of how well the model conforms to common refinement constraint values and these numbers should be close to 1.0. The overall summary report of the model explains that it passed the What check. From observing the data collected from the What check, it can be determined that the homology model contained abnormalities in its backbone conformation as well as unfavorable bond lengths, side chain planarity, and B-factor distribution. However, the only major abnormality would be the backbone confirmation (Table 1). A score of -32.479 suggests that the backbone conformation of the homology model is unusual for the protein. This can be attributed to the backbone not being recognized by the software, and thus not being able to be compared to other backbone structures in the database. To fix this issue, the backbone should be compared with database structures, while taking advantage of C-alpha superpositions with some limitations to the position

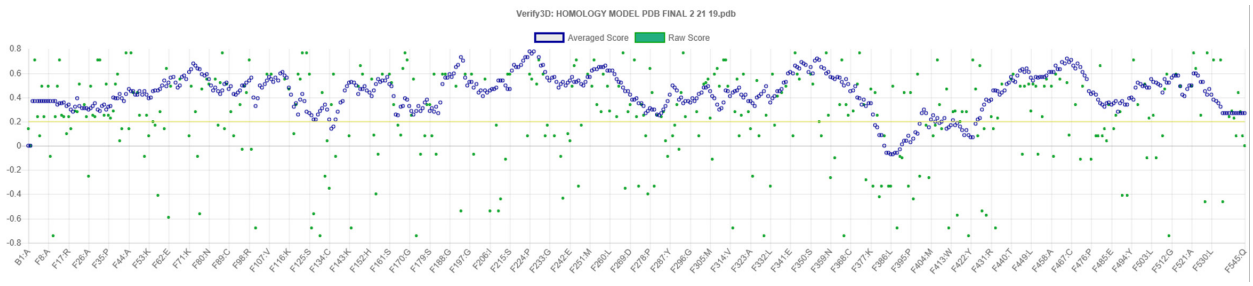


Figure 8. *Verify 3D of Homology model*

Table 1. *What check results for human FAAH homology model.*

Overall Summary Report : PASS	
1st generation packing quality	-0.794
Ramachandran plot appearance	-0.805
Chi-1/Chi-2 rotamer normality	-1.119
Backbone conformation	-32.479
Bond lengths	0.404
Bond angles	0.660
Omega angle restraints	0.909
Side chain planarity	0.262
B-factor distribution	0.330

of oxygen in the backbone. Although, the What check flagged these abnormalities, the values are still within reasonable range to determine that the model is valid.

Knowing that the homology model passed the validation tests, the binding pocket was then identified (in blue) in order to dock compounds for future experiments (Figure 9). ICM Pro docking algorithm

Figure 9. *Identification of binding pocket for human FAAH homology model*

made a box around the predicted ligand binding site and the energy maps of the environment within the docking box were constructed. Using the information from the literature and mutagenesis binding data of known FAAH inhibitors, we visually inspected the binding site in homology model and the location of the binding site is in the agreement with the current literature.

CONCLUSION

Creating a homology model for an unknown crystalized structure can provide valuable insights in structure-based drug discovery and future drug design. Although, the homology model created for this experiment did show some minor abnormalities, it did produce positive results overall and will be used for future docking experiments and will help guide our future medicinal chemistry efforts and studies.

REFERENCES

1. Ahn, K.; Johnson, D. S.; Cravatt, B. F., Fatty acid amide hydrolase as a potential therapeutic target for the treatment of pain and CNS disorders. *Expert Opin Drug Discov* **2009**, 4 (7), 763-784.
2. Wang, X.; Sarris, K.; Kage, K.; Zhang, D.; Brown, S. P.; Kolasa, T.; Surowy, C.; El Kouhen, O. F.; Muchmore, S. W.; Brioni, J. D.; Stewart, A. O., Synthesis and evaluation of benzothiazole-based analogues as novel, potent, and selective fatty acid amide hydrolase inhibitors. *J Med Chem* **2009**, 52 (1), 170-80.
3. Seierstad, M.; Breitenbucher, J. G., Discovery and development of fatty acid amide hydrolase (FAAH) inhibitors. *J Med Chem* **2008**, 51 (23), 7327-43.
4. Otrubova, K.; Boger, D. L., alpha-Ketoheterocycle-based Inhibitors of Fatty Acid Amide Hydrolase (FAAH). *ACS Chem Neurosci* **2012**, 3 (5), 340-348.
5. Otrubova, K.; Ezzili, C.; Boger, D. L., The discovery and development of inhibitors of fatty acid amide hydrolase (FAAH). *Bioorg Med Chem Lett* **2011**, 21 (16), 4674-85.
6. Wheeler, D. L.; Barrett, T.; Benson, D. A.; Bryant, S. H.; Canese, K.; Chetvernin, V.; Church, D. M.; Dicuccio, M.; Edgar, R.; Federhen, S.; Feolo, M.; Geer, L. Y.; Helmberg, W.; Kapustin, Y.; Khovayko, O.; Landsman, D.; Lipman, D. J.; Madden, T. L.; Maglott, D. R.; Miller, V.; Ostell, J.; Pruitt, K. D.; Schuler, G. D.; Shumway, M.; Sequeira, E.; Sherry, S. T.; Sirotkin, K.; Souvorov, A.; Starchenko, G.; Tatusov, R. L.; Tatusova, T. A.; Wagner, L.; Yaschenko, E., Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **2008**, 36 (Database issue), D13-21.
7. Abagyan, R. A.; Batalov, S., Do aligned sequences share the same fold? *J Mol Biol* **1997**, 273 (1), 355-68.
8. Pecic, S.; Makkar, P.; Chaudhary, S.; Reddy, B. V.; Navarro, H. A.; Harding, W. W., Affinity of aporphines for the human 5-HT2A receptor: insights from homology modeling and molecular docking studies. *Bioorg Med Chem* **2010**, 18 (15), 5562-75.

9. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E., The Protein Data Bank. *Nucleic Acids Res* **2000**, 28 (1), 235-42.

10. Li, M.; Wang, B., Homology modeling and examination of the effect of the D92E mutation on the H5N1 nonstructural protein NS1 effector domain. *J Mol Model* **2007**, 13 (12), 1237-44.

11. Morris, A. L.; MacArthur, M. W.; Hutchinson, E. G.; Thornton, J. M., Stereochemical quality of protein structure coordinates. *Proteins* **1992**, 12 (4), 345-64.

12. Colovos, C.; Yeates, T. O., Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci* **1993**, 2 (9), 1511-9.

13. Eisenberg, D.; Luthy, R.; Bowie, J. U., VERIFY3D: assessment of protein models with three-dimensional profiles. *Methods Enzymol* **1997**, 277, 396-404.

14. Pontius, J.; Richelle, J.; Wodak, S. J., Deviations from standard atomic volumes as a quality measure for protein crystal structures. *J Mol Biol* **1996**, 264 (1), 121-36.

Developing a DNA Nano-Robot for Chemotherapeutic Drug Delivery

Suzette Herrera, Allison Serrano, Daniel Arroyo, and Axel Alvarez-Loya
Advisors: [1]Madeline Rasche, Ph.D., and [2]Nina Robson, Ph.D.
[1]*Biochemistry and Chemistry Department, California State University, Fullerton*
[2]*Mechanical Engineering Department, California State University, Fullerton*

ABSTRACT

One of the treatments for cancer is the use of chemotherapy, which involves the administration of anti-cancer drugs. The drugs are given to patients through the bloodstream and radiate throughout the body, attacking the growth of all cells. Although this treatment is effective in fighting cancer, the drugs may have toxicity that also inhibits healthy cells, which produces various side effects. With the advancement of technology, the creation of drug delivery nanostructures made of folded DNA (DNA origami) is now possible. In this work, caDNAno software was used to design an original static circular structure and a bio-nanorobot intended to act as a delivery system for the chemotherapeutic drug doxorubicin. The nanorobot was created using DNA origami which self-assembles into a cancer-targeting nanotube. In initial studies, transmission electron microscopy confirmed the folding of the static design into a circular nanostructure. When analyzed by agarose gel electrophoresis, both the static circular design and the dynamic nanorobot displayed a uniform molecular weight, consistent with successful folding of the structures. This dynamic nanorobot design has potential to promote the prevention of cancer while decreasing side effects using biological material as a specific drug delivery system.

INTRODUCTION

Cancer is the name of a collection of diseases in which an infected cell begins to divide rapidly, growing into surrounding tissues. The cancer cell does not follow normal division pathways; instead, the cell multiplies at a higher rate and begins to build up in any part of the body. Some cancer cells may break off the primary cancer complex and travel throughout the body; this is called metastasis. Most cancer cells receive their nutrients from the blood vessels, creating a direct linkage between the cell complex and vessels surrounding it. This linkage is how the cells can travel throughout the system (Torre et al.2015).

Cancer has a major impact on society today and is the leading cause of death worldwide. In 2012, the National Cancer Society reported 14.1 million new cases of cancer and 8.2 million cancer-related deaths worldwide (Torre et al. 2015). The number of new cases continues to rise every year. Millions of dollars are spent within the cancer research community to discover more insights into how these cells form and how to cure the disease. In modern-day treatment, a cancer patient will undergo various forms of treatment such as surgery and chemotherapy. The treatment will be specific to their disease progression and type. Almost all types of cancers are treatable with chemotherapy, which involves injecting the patient with anticancer drugs that are toxic to the cells, resulting in the killing of cancer cells throughout the body. There are many types of anti-cancer pharmaceuticals available, all having different mechanisms of attacking the cancerous cell. Although they inhibit the growth of cancer in distinctive ways, they are all accompanied by side effects. These side effects include nausea, hair loss, fatigue, loss of appetite, easy bruising. All these side effects take a large toll on the patients' health and strength. Chemotherapy treatment is correlated with side effects because most drugs not only

attack infected cells but healthy cells as well, inhibiting the growth of healthy cells (Torre et al. 2015). However, chemotherapy is the leading treatment along with surgery to overcome the disease.

Doxorubicin ($C_{27}H_{29}NO_{11}$) is a common anticancer chemotherapeutic drug used for solid tumors and acute leukemias (Rivankar, 2014). Doxorubicin is a cell cycle-nonspecific drug which is most active in the S-phase. Doxorubicin will slow or stop the growth of cancer cells by blocking the enzyme Topoisomerase II. This enzyme is needed by the cancer cell to divide and grow. Doxorubicin is currently delivered through the bloodstream allowing access for the drug to attack both healthy and infected cells. Delivery of the drug directly to the infected cell would allow healthy cells to continue to live and potentially eliminate some of the damaging side effects of chemotherapy (Zhang et al. 2014).

Many studies within the past ten years have focused on developing alternative ways to administer anti-cancer drugs specifically to infected cells, allowing the reduction of the side effects correlated with treatment. One of the new systems is bio-nanorobotics. In this method, a nanostructure is designed with biological materials that hold mechanisms allowing the nanorobot to be capable of actuation, intelligence, and sense (Bloh 2013). Bio-nanorobots are made using biological materials such as DNA, proteins, and RNA. DNA molecules are a promising strategy to advance medical technology because they have been shown to be excellent substrates for the construction of dynamic mechanical molecular devices (Li et al. 2018).

DNA origami is a process that enables the rational design and production of DNA nanostructures with controlled size and shape. In this technique, a single-stranded viral DNA strand (M13mp18) (the template or scaffold) self-assembles with staple strands (primers composed of 15-50 bases of nucleotides) during an annealing process to form double-stranded helical DNA structures called nanotubes. These DNA nanotubes are the biomaterial utilized in designing nanoscale DNA structures, since the helices act as a rigid body or a link (Su et al. 2015). In contrast, regions of single-stranded template DNA connecting the helices together will allow movement to occur between the nanotubes, and thus, single-stranded DNA can be considered as joints or connections for these links (Su et al. 2015).

In the current study, using DNA origami, we have designed a nanorobot intended to deliver a chemotherapy drug specifically to cancerous cells. To accomplish this goal, first the origami structure and drug to be delivered were chosen. Second, the DNA origami rectangular shape was modified to allow attachment of the drug. Finally, rectangular shaped DNA nanorobots that self-assemble into tube-like formations were designed to enclose and deliver the drug to cancerous cells.

METHODS

Programming of DNA master thermocycler

An Eppendorf 5331 Mastercycler gradient (Eppendorf, Hauppauge, NY) was used to apply the temperature gradient to form all DNA origami structures. The programmed decrease of 1°C per min from 95-20°C was modified from previously explored protocols (Rothemund 2006).

Design and editing of static structures

M13mp18 viral DNA (Bayou Biolabs, Metairie, LA) was used as the scaffold backbone for DNA nanostructure designs. Circular structures were designed with the software caDNAno, which allows for creation and editing of staple (DNA primer) strands (Douglas et al. 2009). The software option to add staple strands automatically was utilized to create 15-50 bases long staples but required optimization of the segments that bind directly to the scaffold backbone. The binding segments were designed to have a minimum of 7-10 base pairs (bp) on both sides for binding to the scaffold DNA.

Formation of static structures

The thermocycler was programmed to decrease

1°C/min from 95-20°C (Rothemund 2006). M13mp18 scaffold DNA (5 nM) and staple strands (50 nM) were added to a buffer containing 40 mM Tris, 20 mM acetic acid, 2 mM EDTA, and 12.5 mM magnesium acetate, pH 8.0 (TAE-Mg buffer).

Design and Formation of DNA Complex

For creation of the dynamic DNA nanostructure, initial design and optimization were performed using the caDNAno software. This program can be paired with Autodesk Maya (2015) to provide a 3D representation of the nanotubes in the structure (Douglas et al. 2009). Single-stranded M13 viral DNA (10 nM) was mixed with 80 nM short staples and functional strands of DNA in TAE-Mg buffer. The DNA was then incubated in a thermocycler with the same program used for static structures.

Assessment of DNA formation

DNA complex formation was assessed through 1% agarose gel electrophoresis with ethidium bromide staining. Each well held 16 µL of DNA sample and 4 µL of 5X DNA loading dye (Bio-Rad, Hercules, CA). Samples were run at 100 V for 1 hour. Static structures were also visualized using transmission electron microscopy (TEM) in collaboration with Dr. Thomas Bobik at Iowa State University (Ames, IA) and atomic force microscopy (AFM) with Dr. Richard Chang at the University of California, Irvine (Irvine, CA).

Doxorubicin Binding to DNA

Thymine-rich single strand DNA (ssDNA) has been previously shown to bind the anticancer drug doxorubicin (dox) (Chen et al. 2016). In this work, the primer GCTTTTTTTGCTTTTTT was designed and used to bind dox to the DNA origami complex. The ssDNA was mixed with dox at a ratio of 10:1. Natural fluorescence of dox (278 nm) was used to determine successful binding along with the increase in fluorescence between 250-260 nm when the formation of dox-ssDNA complex was formed.

DNA and chemicals

Staples (primers) were synthesized by Integrated DNA Technologies (Coralville, IA). Unless otherwise noted, other chemicals were purchased from Fisher Scientific (Chino, CA).

RESULTS AND DISCUSSION

Static structures and first original DNA origami design

To familiarize the team with the caDNAno DNA design software, an original circular static structure referred to as the "Circle F" design was generated (Figure 1).

A control circular static structure (Rothemund 2006) was also used to optimize the temperature cool down ramp (from 95°C to 20°C) for annealing the template and staple strands for DNA origami structures.

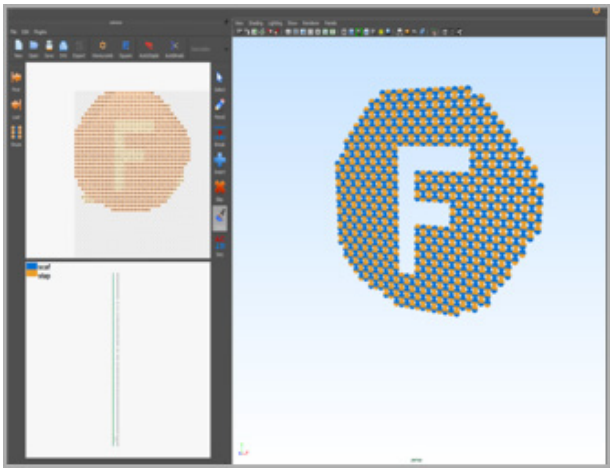


Figure 1. *Computational design of an original static structure (Circle F) using caDNAno software.*

Both the control static circle design and the original Circle F design were tested under three different thermocycling conditions (Figure 2). The programs included (A) a 1°C/min drop over 75 min, (B) a 0.1°C/min decrease over 12.5 hours, and (C) a 0.1°C/6 sec drop over 75 min. The programs averaged 1°C per minute (A and C) and 0.1°C per min (B), as shown in Figure 2. DNA gel electrophoresis allowed the verification of the

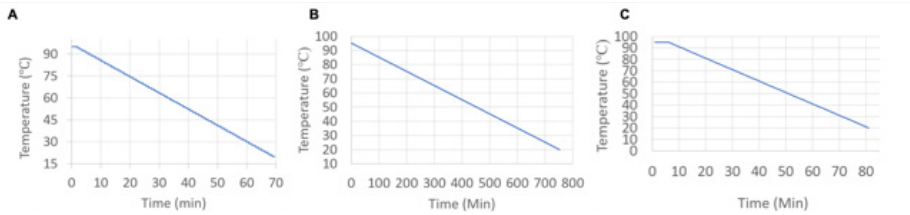


Figure 2. *Mastercycler temperature gradient program. The graphs demonstrate a linear rate of temperature decrease at (A) 1° per min as a step gradient; (B) 0.1° per min; and (C) 1° per min in finer step gradient (0.1°C per 6 sec).*

control and original design reactions at each different annealing condition.

Gel electrophoresis imaging indicated that the thermocycling programs averaging 1°C/minute had a higher efficiency of DNA folding than a program at 0.1°C/minute (Figure 3). With the 1˚C/minute drop, a distinct band between the molecular weight markers of 6000 and 4000 base pairs was observed for the control circular design reaction (Figure 3A, lane 3), indicating the folding of intact nanostructures with a uniform size.

In contrast, the gel images for reactions produced with the 0.1˚C/min program over 12.5 hours had smearing throughout the lanes of both the control circular structure and the original Circle F design (Figure 3B, lanes 3 and 5), indicating DNA structures of various sizes with differing numbers of staple strands attached. Within the smear, only a faint band occurred between the molecular weight of 6,000-4,000 base pairs. Smears below the molecular weight of 500 base pairs represent primers of the static structure that did not bind to the template DNA strand, since the primers are relatively small in molecular weight compared to an assembled nanotube structure. Thus, the program in Figure 2C called SAM02 was used in the construction of subsequent DNA nanostructures.

Atomic force microscopy (AFM) and transmission electron microscopy (TEM) were also used to observe the formation of the static circle control. In both the AFM image (provided by UC Irvine) and the TEM images (courtesy of Iowa State University), the static circle control had a circular shape with a diameter of approximately 200 nm (Figure 4). However, there was a lack of clarity in viewing the holes in the center of the circle that would have created the Circle F images.

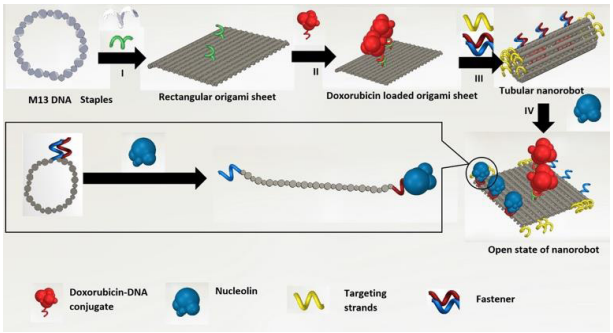


Figure 3. *Design of doxorubicin-DNA nanorobot Illustration of the construction of doxorubicin loaded nanorobot by DNA origami and its adopted function. (I) Single-stranded M13 phage genomic DNA is linked by short staple strands in the formation of a rectangular DNA sheet. (II) Doxorubicin is loaded onto the surface with affinity to capture strands. (III) An addition of fastener strands and targeting strands on the nanorobot to allow for the self-assembly of the nanotube configuration and direction of the nanorobot (IV) the opening of the nanotube after a conformational change between the fastener strands and nucleolin.*

DNA Nanorobot Design

A DNA nanorobot was designed to deliver a chemotherapeutic drug to an active cancer cell. The nanorobot is based on a self-assembled DNA origami nanotube with functional components and attached doxorubicin molecules. A rectangular shaped DNA origami, measuring 90 nm × 60 nm × 2 nm was designed using M13 bacteriophage genome DNA and staple strands through modification of designs by Li et al. (2018). Curvature strands were designed to manipulate the DNA sheet to fold into a nanotube.

For doxorubicin loading, capture strands with poly-GC sequences were extended at four designated

staple strand locations on the surface of the rectangular DNA sheet. The doxorubicin is naturally attracted to poly GC- nucleotides (Chen et al. 2016) and is predicted to anchor itself to the surface of the DNA sheet. Next, fastener strands were designed to close the rectangle into a hollow tube-shaped DNA nanorobot containing doxorubicin molecules. This encloses the drug within the inner tube surface, therefore shielding doxorubicin from attacking healthy cells. To ensure the DNA nanorobot is delivered to the active cancer cell, targeting strands were designed to lead the nanorobot to the surface of the actively reproducing tumor cells. It is hypothesized that the fastener strands recognize the nucleolin proteins on the surface of the tumor cell (Chen et al. 2016) allowing a reconfiguration of the nanorobot to expose the doxorubicin cargo. The cancer cell would then consume the doxorubicin molecules and potentially stop reproducing as doxorubicin inhibits the enzyme Topoisomerase II, which is needed by the cell for division and growth.

Initial Assembly of the DNA Nanorobot

Construction of the rectangular sheet design for the nanorobot was tested using the adapted thermocycling program SAM02, which begins with an initial temperature of 95°C and decreases to 20°C by 1°C per minute using increments of 0.1°C per 6 sec.

When the reaction for the rectangular sheet was analyzed by DNA gel electrophoresis, a single major band appeared between 6,000 to 10,000 base pairs (Figure 5, lane 2), slightly above that of the control static circular structure (Figure 5, lane 5). This result provides initial evidence for the successful construction of nanostructures of uniform size.

Finally, the ssDNA tails designed for anchoring the doxorubicin to the rectangular sheet were added, and the reaction products were compared with the rectangular sheets without tails. DNA gel electrophoresis displayed similar bands for the both the formation of the rectangular sheet without and with the doxorubicin tails attached (Figure 5, lanes 3 and 5, respectively), indicating that the rectangular sheet formation was successfully produced and was not negatively affected by the addition of the doxorubicin tails.

The CSUF team and collaborators are currently modifying the microscopy processes to obtain better clarity in the transmission electron microscopy (TEM) images, which may be improved through decreasing the amount of stain used, utilizing a different type of stain, or providing fresher samples when imaging the DNA nanoscale structures.

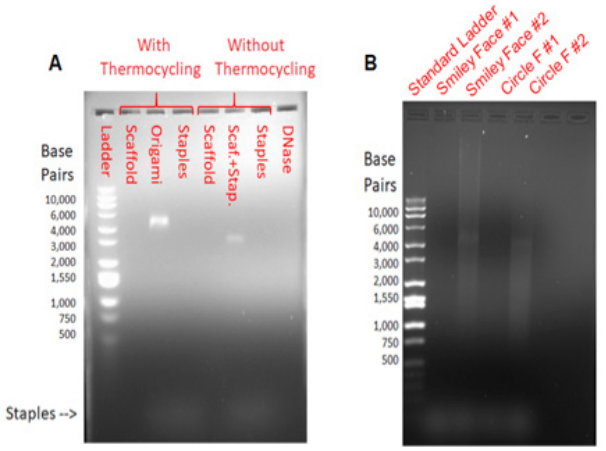


Figure 4. *DNA gel electrophoresis of Origami structures at (A) average 0.1°C per minute and (B) 0.1°C per minute.*

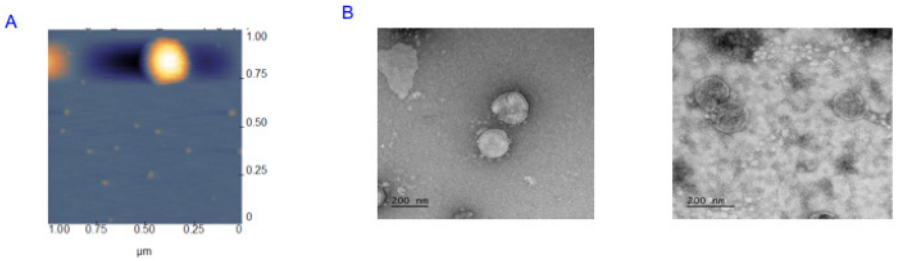


Figure 5. Imaging attempts of the circular structure using (A) AFM microscopy and (b) TEM microscopy

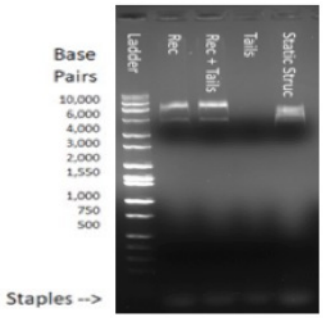


Figure 6. *Gel Electrophoresis demonstrates similar bands indicating uniform folding of DNA origami rectangular sheets with and without Dox-trails.*

CONCLUSION

The CSUF Titan Nano Team has successfully overcome the limitations of the Eppendorf 5331 Mastercycler gradient ramp range to adapt the instruments in the CSUF lab
for DNA origami reactions. Gel electrophoresis and microscopy images confirmed that the static structures retained a circular shaped, which encouraged the team to pursue the creation of biomechanisms with medical implications. We aspire to transition from static to dynamic DNA nanoscale mechanisms for the purpose of targeted cell drug delivery as a new form of treatment for cancer (Li et al., 2018).
The DNA nanorobot in this study was designed to deliver the anticancer drug doxorubicin and holds the potential to contribute to the discipline of nanomedicine.

This field has made remarkable achievements in recent years (Zeng et al.2018) and continues to advance the practice of cancer chemotherapy treatment.

AKNOWLEDGEMENTS

This work was funded by National Science Foundation grant CHE-1508801 and a 2017
CSUPERB Curriculum Development grant. The authors gratefully acknowledge the support of the National Science Foundation NSF award Id #1751770.

We are also grateful to Thomas Bobik and the Iowa State University Microscopy Center for the TEM images and to UC Irvine for the AFM images. We are also grateful to Samuel Fayard for the adapted design of the thermocycler program, Binyun Chen for his illustrations one of which is featured in this article(Figure 3) and Noelle Fitchett for her contributions throughout the applications.

REFERENCES

Bloh A.M. (2013) Nanotechnology and Safety Phar macology. In: Vogel H.G., Maas J., Hock F.J., Mayer D. (eds) Drug Discovery and Evaluation: Safety and Pharmacokinetic Assays. Springer, Berlin, Heidel-berg

Chen, X., Zhou, L., Wang, J., Jiang, G., Cheng, H., & Pei, R. (2016). The Study of the Interaction between Doxorubicin and Single-Stranded DNA. *Chemistryselect,1*(13), 3823-3828. doi: 10.1002/slct.201600473

Douglas, S.M., Marblest|one, A.H., Teerapittayanon, S., Vazquez, A., Church, G.M., Shih, W.M. (2009) Rapid prototyping of 3D DNA-origami shapes with caDNAno. Nucleic Acids Research **37**(15):5001-5006.

National Center for Biotechnology Information. PubChem Compound Database; CID=31703, https://pubchem.ncbi.nlm.nih.gov/compound/31703 (accessed 2019).

Li, S., Jiang, Q., Liu, S., Zhang, Y., Tian, Y., Song, C., Wang, J., Zou, Y., Anderson, G.J., Han, J., Chang, Y., Liu, Y., Zhang, C., Chen, L., Zhou, G., Nie, G., Yan, H., Ding, B., and Zhao, Y. (2018) A DNA nanorobot functions as a cancer therapeutic in response to a molecular trigger in vivo. Nature Biotechnol. **36**:258–264.

Mavroidis, C. (2004). *BIO-NANO-MACHINES FOR SPACE APPLICATIONS*. Presentation, Northeastern University, Boston, Massachusetts.

Rivankar, S. (2014). An overview of doxorubicin formulations in cancer therapy. Journal of Cancer Research and Therapeutics 10(4): 853. doi:10.4103/0973-1482.139267

Rothemund, P.W.K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297–302 (2006)

Su, H., Castro, C., Marras, A., and Zhou, L. (2017). The kinematic principle for designing DNA Origami mechanisms: Challenges and opportunities. Journal of Mechanical Design. 139. 10.1115/1.4036216

Torre, L. A., Bray, F., Siegel, R. L., Ferlay, J., Lortet-Tieulent, J., & Jemal, A. (2015). Global cancer statistics, 2012. CA: A Cancer Journal for Clinicians 65(2): 87-108. doi:10.3322/caac.21262

Zhang, Q., Jiang, Q., Li, N., Dai, L., Liu, Q., & Song, L. et al. (2014). DNA Origami as an In Vivo Drug Delivery Vehicle for Cancer Therapy. *ACS Nano* **8**(7): 6633-6643. doi: 10.1021/nn502058j

Zeng, Y., Liu, J., Yang, S., Liu, W., Xu, L., & Wang, R. (2018). Time-lapse live cell imaging to monitor doxorubicin release from DNA origami nanostructures. *Journal of Materials Chemistry B* **6**(11): 1605-1612. doi: 10.1039/c7tb03223d

On Predictive CD Spectra For Mandelic Acid And Other Small Chiral Organic Molecules, A Computational Approach

Steven Gonzalez Guillen
Advisor: Andrew Petit, Ph.D.
Department of Chemistry and Biochemistry, California State University, Fullerton

INTRODUCTION

Chiral molecules are of great importance given that terrestrial life as humans understand it is known to be predominately "one-handed", or "homochiral". Homochirality is readily observable in the most fundamental of all biopolymers such as DNA, RNA, carbohydrates, and proteins along with their constituent monomers L-amino acids and D-sugars. Notably, L-amino acids serve as building blocks for a plethora of protein structures and D-sugars serve as the structural backbones of DNA and RNA molecules. It is yet to be known for certain the origin of chirality in biochemistry, but there are recent theories that hypothesize that enantiomeric differentiation is due to inter- and circumstellar circularly polarized light (*cpl*). Previously, Bredehöft and coworkers have theorized a model of photochirogenesis and how it could explain the asymmetry found in nature. They theorize that asymmetric photochemistry provides a viable source of enantiomeric excess (e.e.) in precursor molecules in the first living systems. Furthermore, their model predicts that the first asymmetrical photochemical reactions took place in space and were catalyzed by the exposure to circularly polarized light. Bredehöft and coworkers also theorize that photochirogenesis helps explain the dominance of one enantiomeric form found in excess amounts of L-amino acids in carbonaceous chondritic meteorites and ice mantles of inter- and circumstellar dust grains.[1, 2]

Chiral chemistry was first discovered by Louis Pasteur in 1848 when he first single-handedly separated the two isomers of sodium ammonium tartrate. Ever since, chirality has been a major aspect of organic chemistry. Namely, we see the importance of chiral specific synthesis when dealing with pharmaceuticals and other medicinal compounds given the large percentage of drugs that are chiral compounds. At times, the difference between enantiomers is the difference between life and death. Thalidomide is an exemplary model and a cautionary tale of the delicate nature of chiral chemistry. The two enantiomeric structures of thalidomide are showcased in Figure 1. Thalidomide was marketed and sold to hundred of thousands of women as a non-addictive and non-barbiturate sedative that was used both as an anti-emetic and to treat morning sickness; it was also advertised as being completely safe. S-thalidoamide was confirmed in 1961 and 1962 by two independent clinicians, Lenz and McBride, to have readily caused over 10,000 births plagued with defects, abnormalities, and premature deaths over the span of many years. Fortunately, further chemical analysis of the drug allowed for the determination that R-thalidomide was a medically safe sedative while S-thalidomide was the hazardous teratogenic which plagued thousands

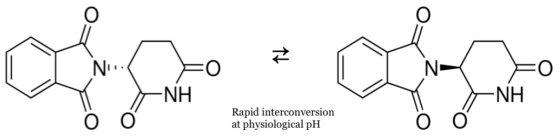


Figure 1. *R-thalidoamide, on the left, is a completely safe sedative that is commonly used by pregnant women to treat motion sickness and morning sickness. S-thalidomide, on the right, is a completely unsafe and biologically catastrophic teratogenic that caused large amount of birth defects, abnormalities, and premature deaths in various countries. Thalidomide is chiral compound which contains a racemic mixture of enantiomers of the same structure. The racemic mixture is approximately 1:1.*

of births. The series of unfortunate events caused by thalidomide were fortunately avoided in the USA between 1957 and 1962. Furthermore, the disaster caused a significant shift in drug testing standards in the USA. The thalidomide disaster further provided evidence of the importance and the possible implications of chiral specific organic synthesis. [3,4,5]

Circular dichroism (CD) spectroscopy (which will be discussed in detail below) has long been used in laboratories to examine and further understand the properties of chiral organic molecules. It is a spectroscopic technique which is used to measure the dichroism of a chiral molecule between the visible (VIS) and ultra-violet (UV) region of the electromagnetic spectrum. For CD spectroscopy to be utilized the molecules being studied must explicitly contain a chiral center (stereo-center). Typical CD studies examine the structures and conformations of larger biomolecules such as carbohydrates, proteins, DNA, and RNA. Additionally, CD studies are completed to further understand how the structures of these molecules are altered by changes in solvent, pH, and temperature. Thus, important information on structure, thermodynamic stability, kinetics, and more can be understood from the CD spectra.

Using circularly polarized light (*cpl*) to drive enantiomeric selective photochemistry presents an alternative approach for selectively synthesizing only one enantiomer from a given enantiomeric pair. In comparison to conventional circular dichroism spectroscopy techniques, synchrotron radiation circular dichroism (SRCD) using *cpl* provides a more suitable experimental method for characterizing the CD spectra for small chiral organic molecules. This is based on the fact that SRCD produces higher intensity light at lower wavelengths, has a higher signal-to-noise ratio, and allows for a faster acquirement of spectral data in comparison to conventional bench top UV-VIS light sources.1 The characteristic photochemical behavior that results upon exposure to *cpl* can serve as a novel alternative to characterizing the spectral differences between R and *S* enantiomers. Moreover, the chiral molecules studied are characterized by their optical rotation once exposed to *cpl*, with the degree of rotation dependent on variables such as enantiomeric purity of the sample, concentration, and the path-length of the spectroscopic cell utilized during SRCD. Thus, the apparent need for enantiomeric selective chemistry in various fields of study can immediately prosper from the characterizations of these small chiral molecules using a computational model to first predict the wavelength at which each chiral molecule will absorb *cpl* to drive enantiomeric selective photochemistry.

The objective of the joint theory-experimental study is to determine the extent to which computational approaches can create a predictive tool for describing the CD spectra of chiral organic molecules. For this family of compounds, Dr. Evans and coworkers have used SRCD to characterize their CD spectra. As shown below, our results demonstrate a generally good agreement between the calculated and CD spectra. As such, the study demonstrates that CD spectra calculated using standard electronic structure approaches are capable of predicting the wavelength ranges in which electronic transitions will occur for such compounds.

Circular Dichroism, CD Spectroscopy, and More

Circular dichroism (CD) is the difference between absorptions when using left circularly polarized light (*lcpl*) and right circularly polarized light (*rcpl*). CD spectroscopy has long been used to study an array of molecules in solid, gaseous, and aqueous phases. For CD spectroscopy to be utilized the molecules being studied must explicitly contain a chiral center (stereo-center), a molecule containing at least one atom with a *minimum* of three different attached atoms or groups where the transposition of two of these attachments will generate another stereoisomer. These optical isomers are referred to as R and *S* enantiomers, *which are mirror images that are non-superimposable*.

With CD or SRCD spectroscopy, a plane of polarized light is split into both *lcpl* and *rcpl*. The

resulting light sources will be elliptically polarized and absorbed to a much greater extent by one enantiomer. Thus, the CD instrument will detect the two different polarized lights separately. Furthermore, the CD spectrum is obtained when the dichroism is measured as a function of wavelength. An example of the CD spectra of a small organic chiral molecule is provided in Figure 2. Moreover, typical CD instruments are used to describe the differential absorption of the enantiomers in the VIS-UV range. The units of measurement are described in terms of the difference in absorption (ΔA), where A is a dimensionless unit representing the absorption intensity. The graphical illustration of absorption in a CD spectra is indicated by having an either positive or negative electronic transition band (or "peak"); this is determined by Equation 1. Thus, a positive peak is correlated to a strong absorption of *lcpl* and a negative peak is correlated to a strong absorption of *rcpl*. The difference in absorption is denoted in Equation 1:[1,6,7]

$$\Delta A = A_{lcpl} - A_{rcpl} \quad (1)$$

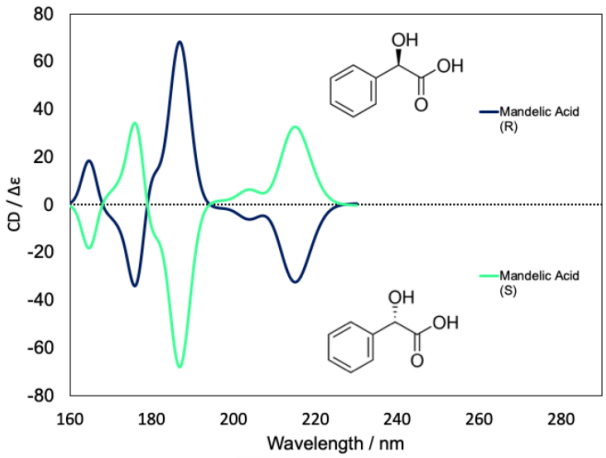


Figure 2. *An example CD spectra of mandelic acid, where the blue spectral lines are the results of the R enantiomer, and the teal spectral lines are the results of the S enantiomer. A typical CD spectra such as this will showcase electronic transition bands or "peaks" that mirror each other identically. Note that the peaks are completely identical to each other to within a minus sign. Peaks above the x-axis (x = 0) absorb lcpl more strongly while peaks below the x-axis absorb rcpl more strongly (see Equation 1).*

It is important to note that CD spectra are sensitive to the solvent in which they are measured. The effects of solvation primarily cause shifts in the position of the absorption bands due to changes to the structural conformations of the chromophore and surrounding solvent molecules as well as changes to the arrangement of partial charges around the chromophore. The ability of a solvent to stabilize the partial charges of solute is parameterized by the dielectric constant, ϵ. In the experiments performed by Dr. Evans and coworkers, both water and 1,1,1,3,3,3-hexafluoro-propan-2-ol (hexaflouroisopropanol, HFiP) were used; $\epsilon_{water} = 78.30$ and $\epsilon_{HFiP} = 16.70$. During our computational analysis we used ϵwater = 78.35 and $\epsilon_{2\text{-methyl-1-propanol}} = 16.77$; 2-methyl-1-propanol was a computational analogue for HFiP as the Gaussian software does not contain solvent parameters for HFiP. The structures of the solvents are showcased in Figure 3. It should be noted that water's increased dielectric constant is indicative of its increased polarity and its ability to more readily insulate electrical charges from one another and thus more strongly stabilize solute partial charges.[2, 7]

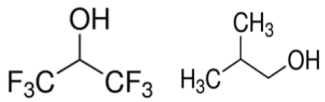


Figure 3. *The molecular structures of the solvents used for the longevity of the joint theoryexperimental study. Water has* $\epsilon_{water} = 78.30$, *and* $\epsilon_{water} = 78.35$ *computationally. HFiP has* $\epsilon_{HFiP} = 16.70$. *2-methyl-1-propanol was used as a computational analogue for HFiP, has* $\epsilon_{2\text{-methyl-1-propanol}} = 16.77$.

Previous studies have discussed what experimental factors affect the effectiveness of enantiomeric selective photochemistry. Specifically, the enantiomeric excess (*e.e.*) which is typically denoted as a percentage (*e.e.*%), quantifies the difference in the relative yields of the *R* and *S* enantiomers after a reaction. Notably, the maximal achievable *e.e.* is not exclusively dependent on the differential absorption, ΔA, obtained from CD spectroscopy. As seen in Equation 3, maximal achievable

e.e. is also dependent upon the extinction coefficient, ε, of the chiral molecule. The extinction coefficient, also referred to as molar absorptivity or molar extinction coefficient, is the degree to which a given solution will absorb light; where an increase in ε is an increase in light absorbed per mole of solute. As seen in Equation 3, calculating *e.e.* is further reliant upon the anisotropy factor, *g*, which is the ratio of the CD spectrum to the UV-VIS absorption spectrum. The anisotropy factor is directly indicative of the predicted maximal *e.e.* achievable by asymmetric photolysis using *cpl* to a given extent of reaction, ξ, ($0 \leq \xi \leq 1$). In the context of photochemistry, the extent of reaction reflects the fraction of molecules which undergo reaction after absorbing a photon.[1,2]

$$g = \frac{\Delta A}{\frac{1}{2}\varepsilon_{rcpl} + \frac{1}{2}\varepsilon_{lcpl}} \tag{2}$$

$$e.e. \geq (1 - (1 - \xi)^{\frac{g}{2}}) * 100\% \tag{3}$$

The Importance of Model Systems & Building Blocks

The main focus of this proof of principle study are the following molecules: mandelic acid ($C_8H_8O_3$), benzoin ($C_{14}H_{12}O_2$), 2-phenylpropionic acid ($C_9H_{10}O_2$), and methyl mandelate ($C_9H_{10}O_3$). These molecules were selected because they are each small chiral compounds that contain a carboxylic acid group. The only exception being benzoin which contains a benzoyl group. This family of small chiral molecules were further selected given their innate ability to polymerize into small

$$G_k = E_k + H_k - TS_k \tag{4}$$

biodegradable building blocks (also referred to as biodegradable polymers). Building block molecules are of particular interest given their versatility in organic synthesis, pharmaceuticals, and materials. Buchard and coworkers pointed out that mandelic acid can be polymerized into a highly isotactic type of biodegradable plastic referred to as PMA (poly-mandelic acid). PMA is a more eco-friendly analogue to polystyrene and other essential plastic products used by everyday consumers. Thus, by creating a model to predict the CD spectrum of building block molecules we can determine the particular wavelength at which molecules such as mandelic acid absorb *cpl*. Therefore, researchers such as Buchard can benefit from having a refined understanding of how to polymerize biodegradable plastic products using *cpl* as a greener alternative for enantiomeric selective photochemistry.[8]

Computational Methods and Computational Chemistry

Optimized geometries for all conformations of the molecules of interest were performed using density functional theory (DFT) with the B3LYP functional and 6-31+G** basis. Geometry optimization were performed on multiple molecular conformations of each molecule in order to sample thermal equilibrium. The optimized molecular geometries were verified to be minima by calculating the vibrational frequencies. Water was the default solvent used to match the experimental conditions ($\epsilon_{water} = 78.35$). Experiments on two molecules, mandelic acid and benzoin, were performed in HFiP ($\epsilon_{HFiP} = 16.70$). Calculations performed on these compounds used 2-methyl-1-propanol ($\epsilon_{2\text{-methyl-1-propanol}} = 16.77$) as a computational analogue to HFiP in order to have a comparable dielectric constant value.[9] The free energy of each conformation was then calculated using Equation 4, where E_k is the electronic energy of the kth conformation , H_k is the enthalpy, S_K is the entropy, and T is the temperature.

$$< CD > = \sum_k CD_k \frac{exp(-G_k/RT)}{q} \tag{5}$$

At each optimized molecular geometry, the CD spectrum was calculated using time dependent density functional theory (TDDFT) with the ωB97XD functional and the 6-311++G** basis, and a polarizable continuum solvation model with the appropriate dielectric constant. The theoretical line spectra was convoluted with a gaussian line shape with a full-width at half maximum

of 0.1 eV (806.554 cm^{-1}) for all spectra to approximately account for homogenous broadening. The collection of CD data per molecule was then thermally averaged over all unique conformation using Equation 5.

$$q = \sum_{k} exp(-G_k/RT) \tag{6}$$

Here, CD_k is the CD spectrum associated with the kth molecular conformation, G_k is the relative free energy of the *k*th conformation, *R* is the ideal gas constant, and *q* is defined in Equation 6. Note that only conformational species with thermal populations ≥1.0% were used to create the final CD spectra, this included both neutral and anionic species (if applicable, based on pK_a; see Table 1). Furthermore, the CD spectra were graphed using strictly the data collected for the R enantiomer because R and *S* enantiomers will intrinsically produce identical CD spectra to within a minus sign.[7]

Table 1. *A table showcasing the pK_a values of all the molecules studied. Note the increase pK_a as the attachments to the chiral carbon increase in complexity.*

Compound	pKa
mandelic acid	3.41
benzoin	12.60
2-phenylpropionic acid	4.59
methyl mandelate	12.19

As mentioned before the inclusion of both neutral and anionic species were solely based on the experimental pK_a, the quantitive measurement of acidity, of each molecule. The use of pKa allowed us to determine the relative weights of the neutral and anionic forms of the molecules. For the molecules studied, the relative weight of neutral species were typically ≥98.0% and ≤2.0% for anionic species. Thus, the final CD spectrum of a given molecule in solvent was influenced to a greater extent by the neutral species and less by the anionic species.

RESULTS AND DISCUSSION

The resulting CD spectra for each respective molecule are showcased here. Note that mandelic acid was studied twice, once in water and once in 2-methyl-1-propanol. Note that ΔA is referred to as Δε in the experimental CD spectra. In Figures 4 - 8, the computational CD spectra is showcased on the left and the experimental CD spectra is showcased on the right.

Mandelic Acid

In Figure 4, the CD spectra of mandelic acid in water shows good agreement with the experimental spectra. Computationally, there are minor peaks located at lower wavelengths at 164.6 nm and 175.8 nm. These

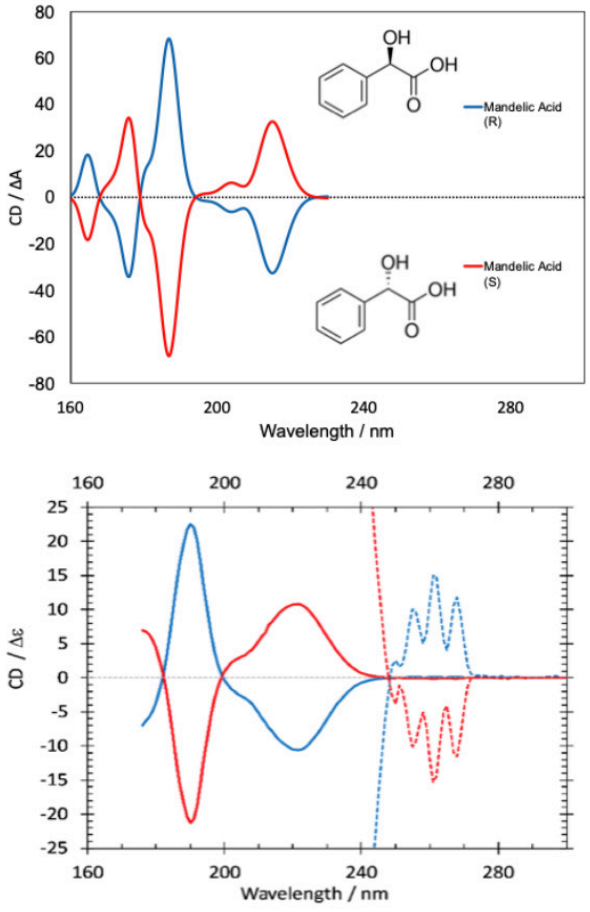


Figure 4. *A comparison between the computational and experimental CD spectra of mandelic acid. The solvent used is water.*

peaks are found experimentally near 175 nm. Next, a major peak is located at 186.6 nm in the computational spectrum while the corresponding peak is at at 190 nm in the experimental spectrum. This peak is attributed to a n → π* transition. In the computational spectrum, this peak includes a small shoulder on its left-hand side which is not resolved experimentally. This is most likely due to the ad-hoc broadening parameter used computationally (0.1 eV or 806.554 cm-1) not exactly matching the homogenous broadening observed in the experimental spectrum. Another set of more broadly spaced peaks were located computationally at 203.8 nm and 215 nm. These are comparable to the peaks found experimentally at 205 nm and 220 nm. Again, the peak found at 220 nm is the result of a n → π* transition. It should be noted that the experimental spectrum is amplified x10 at higher wavelengths (>240 nm) to highlight a series of low-energy spectral features. The computational approach was unable to capture these low-energy features.

In Figure 5, the computational CD spectrum of mandelic acid in 2-methyl-1-propanol shows good agreement with the experimental spectrum in HFiP. The CD spectra of mandelic acid in water and HFiP are similar in terms of their overall shape and the number of peaks. The CD spectrum in HFiP does show a solvatochromic blue shift. It should noted that the experimental spectrum is amplified x10 at higher wavelengths (>240 nm) to highlight a series of low-energy spectral features. The computational approach was unable to capture these low-energy features.

In Figure 6, the calculated CD spectrum of methyl mandelate in water shows good agreement with the experimental spectrum. The only structural difference from mandelic acid is the replacement of a hydroxyl group with a methoxy group on the α-carbon (chiral carbon). As such, the CD spectrum of methyl mandelate is qualitatively similar to that of mandelic acid. Computationally, there are no minor peaks located below 170 nm but there are peaks found experimentally near 175 nm; this will be improved in the future by

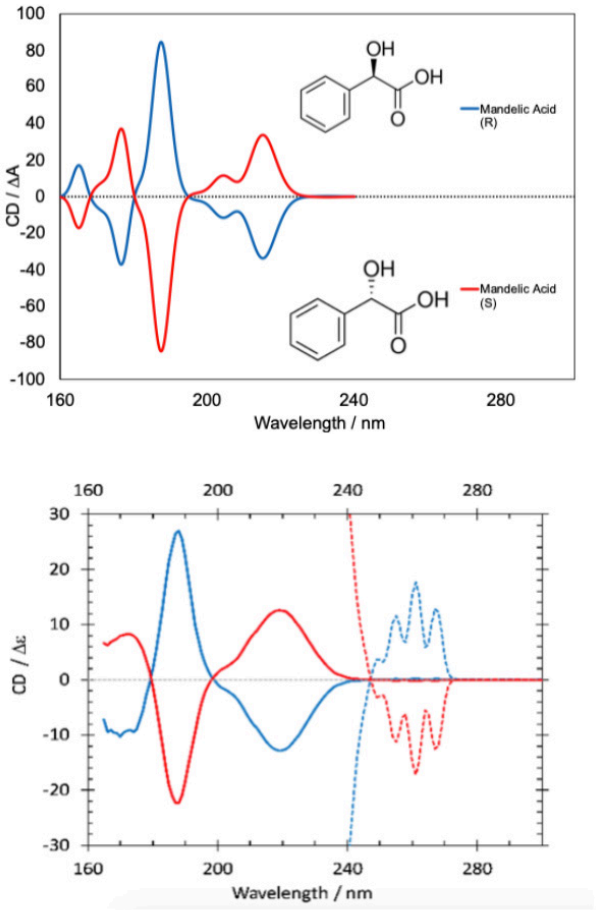


Figure 5. *A comparison between the experimental CD spectrum of mandelic acid in HFiP and the computational CD spectrum in 2-methyl-1-propanol.*

increasing the number excited electronic states included in the calculation. Next, a major peak is located at 186 nm. This major peak is found experimentally at 190 nm and is attributed to a n → π* transition. Another set of more broadly spaced peaks are located computationally at 203.4 nm and 212.4 nm. These are comparable to the peaks found experimentally at 206 nm and 220 nm. Again, the peak found at 220 nm is the result of n → π* transition. It should be noted that the experimental spectrum is amplified x10 at higher wavelengths (>240 nm) to highlight low-energy spectral features. The computational approach was unable to capture these features.

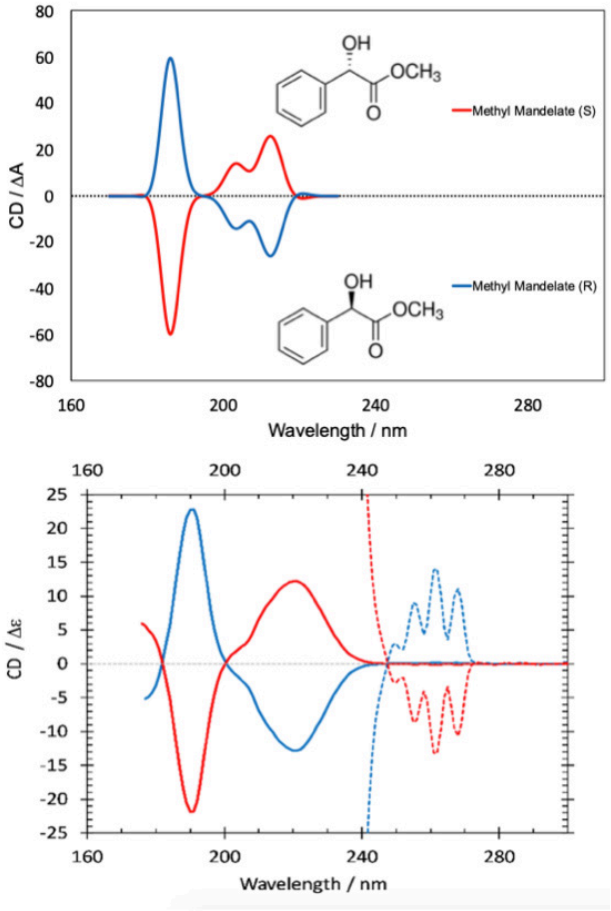


Figure 6. *A comparison between the computational and experimental CD spectra of methyl mandelate.*

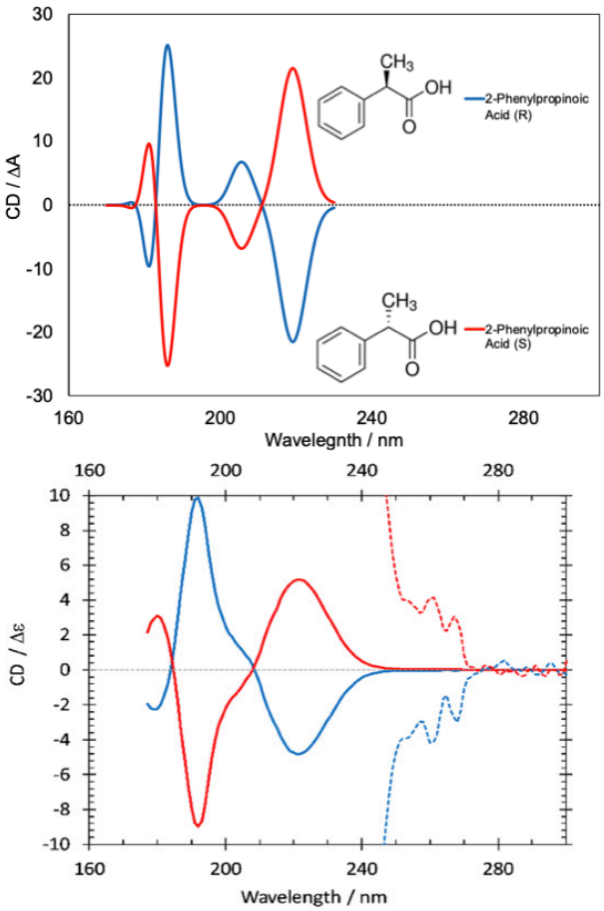


Figure 7. *A comparison between the computational and experimental CD spectra for 2-phenylpropionic acid. The solvent is water.*

In figure 7, the calculated CD spectrum of 2-phenylpropionic acid in water shows good agreement with the experimental spectrum. The only structural difference from mandelic acid is the replacement of a hydroxyl group with a methyl group on the ß-carbon. As such, the CD spectrum of 2-phenylpropionic acid is similar to that of mandelic acid. Computationally, there are no minor peaks located below 170 nm but there is a peak located at 181.2 nm. Next, a major peak is located at 186 nm with a separated shoulder at 205.6 nm. This major peak is located experimentally at 191 nm with a resolved shoulder near 200 nm. This major peak at 191 nm is attributed to a n → π* transition. Next, a major peak is located at 219 nm and is comparable to the peaks found experimentally at 220 nm. Again, the peak found at 220 nm is the result of n → π* transition. It should noted that the experimental spectrum is amplified x10 at higher wavelengths (>240 nm) to highlight these low-energy spectral features.

The least amount of agreement found between the computational and experimental spectra is showcased in Figure 8 for benzoin. Benzoin is unique in this study because it does not contain a carboxyl acid group, rather it contains a benzoyl group. Computationally, there is a set of minor peaks located at 162.4 nm, 166.4 nm, 171.8 nm, and 176.8 nm. These are comparable to those found experimentally at 164 nm, 170 nm, and 178 nm.

The peaks located near 178 nm are likely due to a π → π* transition. Next, a major peak was located at 182.4 nm. This major peak includes a shoulder on its right-hand side near 188.4 nm. This major peak is found experimentally at 189 nm and is attributed to a n → π* transition. Irregularities and less comparable agreement between the two methods are found at wavelengths ≥195 nm. The greatest amount of discrepancy is found at the most intense peak located at 195.2 nm; the experimental peak located at 205 nm is much less intense. Next, two minor peaks are located at 206.2 nm and 226.6 nm; these two peaks do not cross the x-axis and are not broad. Experimentally two minor peaks were located at 215 nm and 232 nm; these two peaks do cross the x-axis and are broad. For benzoin, the computational approach is able to capture spectral data at higher wavelengths (≥240 nm); two minor peaks are located at 248.8 nm and 258.4 nm. Experimentally only one major peak is located in this range at 256 nm and this peak is noticeably broad.

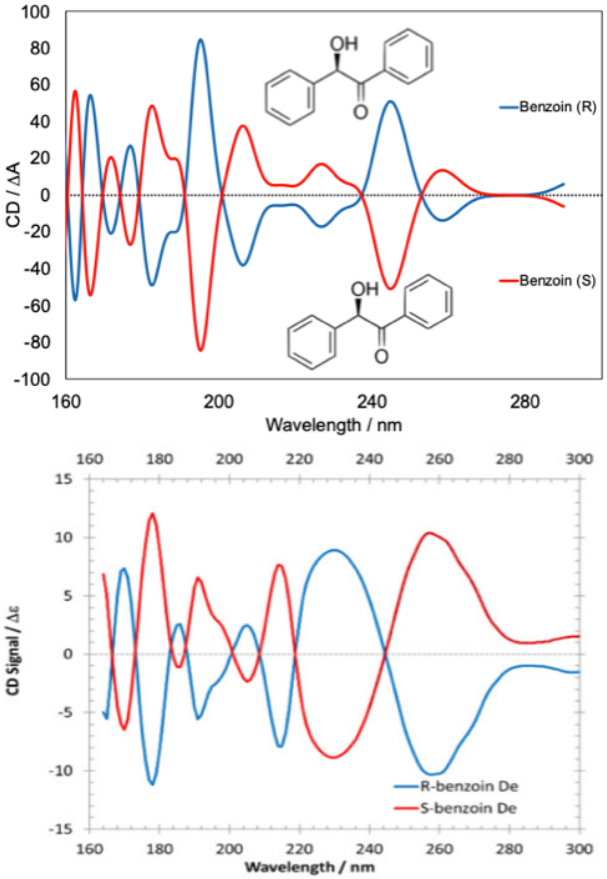


Figure 8. *A comparison between the computational and experimental CD spectra of benzoin. The solvent in the calculations is 2-methyl-1-propanol, an analogue solvent comparable to the solvent HFiP used in the experiments.*

Future work will involve attempting to improve the agreement between the computational and experimental CD spectra of benzoin. The current difficulty when modeling benzoin may arise from a few elements. The first is that the electronic structure of benzoin may less well captured by the current computational approaches than the other compounds, perhaps because of its more extensive pi-conjugation . Secondly, the assortment of molecular conformations may not include the most thermodynamically relevant conformation(s) in the sample pool; the additional benzene ring increases the conformation space that must be searched. Thus, the continued search for the most thermodynamically stable conformation(s) may allow for better agreement between experiment and theory.

CONCLUSIONS

In sum, the computational approach is both a time and cost effective method for predicting the CD spectra of small organic chiral molecules such as those studied here. This study demonstrates the effectiveness of characterizing the CD spectra of small chiral molecules using both computational and experimental approaches in tandem. We are able to provide good agreement when compared to previous experiments using SRCD. Further work should be completed in order to understand and correct the deficiencies observed in the calculated CD spectrum of benzoin in order to create more accurate model which can be widely used for a wide range of small organic chiral molecules. Further work should also increasingly focus on developing an approach for utilizing computational chemistry to predict the anisotropy factor and hence *e.e.* as a function of wavelength. This, along with further experimental developments, will allow for the development of chiral specific photochemistry.

REFERENCES

1. Evans, A. C.; Meinert, C.; Bredehöft, J. H.; Girl, C.; Jones, N. C.; Hoffmann, S. V.; Meierhenrich, U. J. Anisotropy Spectra for Enantiomeric Differentiation of Bimolecular Building Blocks. Topics in Current Chemistry. https://link.springer.com/content/pdf/10.1007/128_2013_442.pdf

2. Bredehöft, J. H.; Jones, N. C.; Meinert, C.; Evans, A. C.; Girl, C.; Hoffmann, S. V.; Meierhenrich, U. J. Understanding Photochirogenesis: Solvent Effects On Circular Dichroism and Anisotropy Spectroscopy. Chirality. 2014, 26, 373-378. https://doi.org.10.1002/chir.22329

3. Nguyen, L. A.; He, H.; Pham-Huy, C.; Chiral Drugs: An Overview. International Journal of Biomedical Science. 2006, 2, 2, 85-100

4. McBride, W. G.; Thalidomide and congenital abnormalities. The James Lind Library. 1961, 2, 1358. http://www.jameslindlibrary.org/mcbride-wg-1961/

5. Lenz, W.; Knapp, K.; Foetal Malformations Due To Thaliomide. *Problems of Birth Defects*. Springer, Dordrecht. T. V. N. Persaud, 1977. https://doi.org/10.1007/978-94-011-6621-8

6. Yao, H.; Wynendaele E.; Xu, X.; Kosgei, A.; De Spiegeleer, B. Circular Dichroism In Functional Quality Evaluation On Medicines. Journal of Pharmaceutical and Biomedical Analysis. 2018, 147, 50-64. https://doi.org.10.1016/j.jpba.2017.08.031

7. Ranjbar, B.; Gill, P.; Circular Dichroism Techniques: Biomolecular and Nanostructural Analyses - A Review. Chemical Biology & Drug Design. 2009; 74, 101-120. https://doi.org.10.1111/j.1747-0285.2009.00847.x

8. Buchard, A.; Carbery D.R.; Davidson, M.G.; Ivanova, P.K.; Jefferey B.J.; Kociok-Köhn, G.I.; Lowe, J.P. Preparation Of Stereoregular Isotactic Poly(mandelc acid) Through Organocatalytic Ring-opening Polymerization Of A Cyclic *O*-carboxyanhydride. Angewandte Chemie. 2014, 53, 50, 13858-13861. https://doi.org.10.1002/anie.201407525

9. *Q-Chem User's Manual, Version 5.0*. Q-Chem, Inc. California, 2017.

10. Warnke, I.; Furche, F.; Circular Dichroism: Electronic. WIREs Computational Molecular Science. 2012, 2, 160-166. https://doi.org.10.1002/wcms.55

11. Atkins, P.; De Paula, J. *Physical Chemistry: Thermodynamics, Structure, and Change, Tenth Edition*; W.H. Freeman and Company: New York, 2014

Using U-Pb Dating to Determine the Maximum Depositional Ages and Provenance of the Cow Creek and Fairview Metasedimentary Pendants, southern Sierra Nevada Mountains

Christian Concha
Advisor: Diane Clemens-Knott, Ph.D.
Department of Geological Sciences, California State University, Fullerton

INTRODUCTION

This study pertains to the Sierra Nevada Mountain range and the rock types contained within, primarily the uplifted and eroded Sierra Nevada Mountain range's exposed Mesozoic granodioritic rocks that crystallized beneath the Sierra Nevada volcanic arc. The Sierra Nevada volcanic arc formed during the Mesozoic Era as the Farallon plate subducted eastward under the North American plate. Magmas from the arc intruded into preexisting rocks, heating them and recrystallizing them. Metamorphism, coupled with deformation accompanying arc emplacement, make it almost impossible to determine the age and origin of pre-Mesozoic sedimentary wall rocks of the arc using fossils, mineralogic and/or structural relations. This leaves much of the history of pre-Mesozoic California unknown.

Despite the difficulty in understanding pre-Mesozoic California, a new analytical strategy provides a way to determine the depositional age and provenance of metasedimentary rocks. U-Pb dating of detrital zircons separated from clastic sedimentary rocks provide an estimate of the maximum depositional age of the unmetamorphosed protolith and the foundation for describing the sediment provenance. A constraint on the maximum depositional age (or MDA) of a sedimentary deposit is estimated from the youngest population of zircon grains in the sediment. The sediment deposit cannot be any older than the maximum depositional age because a sedimentary deposit cannot contain zircons that have not yet formed. Provenance refers to the collection of source rocks that were eroded to make the clastic sediment (e.g., the Appalachian Mountains are the source region of ~1.1 Ga detrital zircons). Pendants, separated from each other by Mesozoic plutons, may be correlated if they have similar provenance and MDA.

Once the zircon grains have been dated, they will be organized into an age-probability plot. These plots are histograms of the U-Pb ages of detrital-zircon samples. Different pendants will be compared are compared by calculating the degree of overlap between curves (a measure of overlapping ages between two samples) and degree of similarity (the proportion of similar ages between samples).

By dating detrital zircon grains separated from multiple pendants, it may be possible to locate pre-Mesozoic fault zones within the Sierra Nevada Mountains, which could lead to a better understanding of western North American tectonics. For example, others have hypothesized that a San Andreas-like transform fault transported PreCambrian-Paleozoic passive margin sedimentary rocks from eastern CA to Sonora, Mexico during the Permian Period. This hypothetical fault also may have transported Paleozoic ocean floor rocks (now called the Kings-Kaweah ophiolite belt) against what is now CA, where they accreted. The present study will provide new data to test and refine these tectonic hypotheses.

BACKGROUND

The rifting of Rodinia was powered by varying magma pulses, one at ~665 to 650 Ma and again at ~500 to 485 Ma (Lund et, al. 2010). This formed the western margin

of Laurentia (Lund et al., 2010). After ~ 200 Ma of rifting and igneous activity, deposits of the Cordillera miogeocline formed along the rift zone as sediments were deposited on top of the rifted margin of the crystalline Mojave province. At approximately 260 Ma, the Kings-Kaweah ophiolite was docked along the Mojave province, and a terrane block of miogeoclinal rocks were later translated south to Sonora, Mexico. Being that the Panthalassan lithosphere (preserved as the accreted King-Kaweah ophiolite) was much colder and denser material than the North American plate, it foundered and began to subduct eastward under the rift margin. This is the point when arc volcanism began to form the Sierra Nevada Batholith.

Deposits of the Osgood Mountains (Nevada) showcase detrital zircon differences in provenance along the Cordilleran miogeocline. Detrital zircon analysis from Lind et al. (2014) demonstrate age differences between several sampling sites from the Osgood Mountain Quartzite and the Preble Formation. Both the Lund and Lind papers showcase that a majority of areas along the Cordillera are metamorphosed sedimentary rocks that were deposited along the passive margin of western Laurentia after the rifting of Rodinia. Ross (1987) described several metasedimentary pendants that are exposed in the southern Sierra Nevada mountain range, and identified locations of quartzite. Most importantly, this work mentioned the presence of detrital zircons in several pendants that could be used for statistical analysis to determine the pendants MDA and provenance.

Ross (1987) described Fairview pendant as mostly quartzite with a lithology that is dark, granular to pebbly, virtually unsorted quartzite. Ross (1987) also described the Cow Creek Pendant as having sugary quartzofeldspathic schist with other varying schists, impure quartzite, and some amphibolite layers along with dark color metasedimentary rocks and metavolcanics rocks.

METHODS

During the summer of 2017, ~3-gallon samples of quartz-rich rocks were collected from the Fairview and Cow Creek pendants. At CSUF, the samples were crushed using a jaw crusher and disk mill. Next, the samples were sieved with a RoTap machine in order to produce a separate of grain-sized particles. This fine-grained sand were then processed using a Wilfley table. The Wilfley uses water with a shaking motion in order to separate grains into "lights" and "heavies", much like gold panning. "Lights" refer to grains that float away into one container while "heavies" refer to grains that sink and are shaken towards the far end of the table into a separate container. The Wilfley heavies were dried and allowed to settle in lithium metatungstate (LMT). Spinning in a centrifuge increased the settling rate. The material that sank in the centrifuge was then passed through a Franz Magnetic Separator to remove any magnetic material. Purification of the zircon-rich separate by hand-picking under a microscope was conducted, as necessary. Finally, the zircon separates were sent to the Arizona Laser Chronology facility where they were mounted in epoxy, polished, and imaged using a scanning electron microscope prior to our arrival in April 2018. My goal was to date approximately 300 zircons per sample.

U-Pb isotopic ratios were measured using an LA-ICP-MS (Laser Ablation Inductively Coupled Plasma Mass Spectrometry). This analytical technology allows direct analysis of highly sensitive elements and isotopes on solid samples. The LA-ICP-MS starts with the ablation (the stripping of finer particles) of the sample using a laser at a desired spot in a crystal. Then the LAICP-MS uses a plasma source to strip off electrons forming charged ions, which are focused though a set of charged plates and accelerated through an electric sector magnet. Ions are separated based on their mass to charge ratio, and are counted in a series of electronic collectors. When collected, they are converted into voltages and compared across different collectors. The LA-ICP-MS software uses the measured isotopic ratios to

calculate a crystallization date for each analyzed zircon grain. Detrital analysis is then conducted in Isoplot using Pb/U Concordia diagrams and relative age probability diagrams. Probability diagrams
show age and uncertainty related to each age. Igneous samples are also plotted in Isoplot using Concordia and weighted mean diagrams. Ages, along with related uncertainties, will also be shown.

Various discordant filters were applied to the Concordia diagrams to filter out discordant grains (i.e. zircon for which the measured 207Pb/235U and 206Pb/238U ratios yield significantly different ages). The industry standard for these filters is usually 20%, but for this study a 70% and a 5% filter was used. The 70% filter is used to generate a wider data pool in order to see the discordance of certain grains within a population. The 5% filter was then applied to determine the unaltered peak ages as defined by the least discordant grains. The youngests population of three or more grains was used to determine the maximum depositional age of the sedimentary protolith.

RESULTS

Results for BFD-105 and BFD-107a are tabulated in Appendix 1 and plotted in Figures 1a- 2f. All histograms display the U-Pb ages as blue bars while the red smooth curves include the errors associated with each age analysis. For BFD-105, major peaks occur at 1.14, 1.42, 1.77, and 2.68 billion years old under a 70% concordance filter (Figure 1a). There is also a small 2.25 (n=3) billion year old peak present in BFD-105. For BFD-107a, major peaks appear at 1.46, 1.78, and 2.61 billion years old under the 70% discordance filter along with a small 2.14 (n=5) billionyear- old peak (Figure 2a).

INTERPRETATIONS

Based on the results, it is safe to say that both the Fairview and Cow Creek pendants are not the same. BFD-107a clearly has lead-loss trajectories that point to major peaks at 2.61, 1.78, and 1.46 billion years old under a 70% discordance filter (Figure 2b, 2c). The most confident maximum depositional age is at 1.46 billion years ago, though some small possible discordant populations appear at 789 Ma (n=3) which indicates pre-passive margin, and 549 Ma (n=1). BFD-105 at 70% discordance shows the presence of a 1.14 Ga age population not present in BFD-107a (Figures 1b, 1c). Along with the 1.14 population, BFD-105 has populations at 1.14, 1.42, 1.77, and 2.68 billion years old which are similar to populations present in BFD-107a. BFD- 105 has a small 900 Ma peak (n=1) though the maximum depositional age is more likely to be at 1.14 Ga due to the amount of grains present at each population. Both samples have a small population

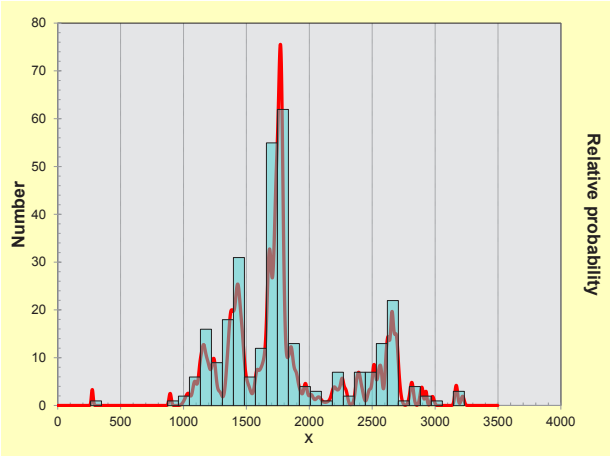


Figure 1a. *Histogram of sample BFD-105 under a 70% discordance filter.*

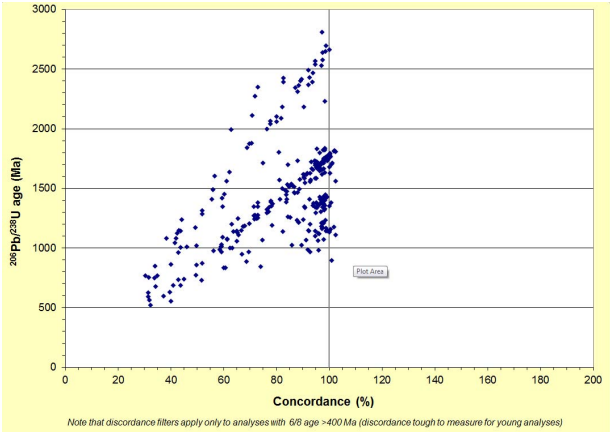


Figure 1b. *Plot of age vs. concordance for sample BFD-105 under a 70% discordance filter.*

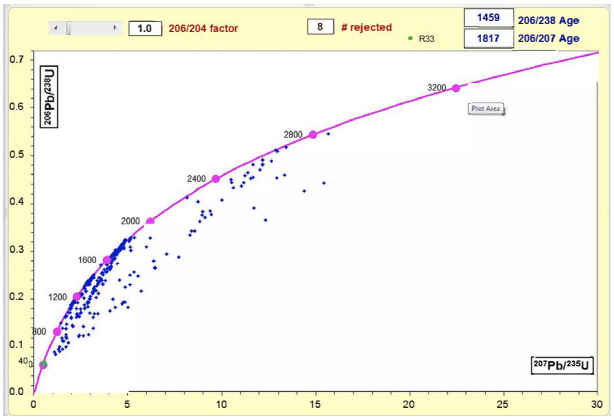


Figure 1c. *Concordia plot under a 70% discordance filter for sample BFD-105.*

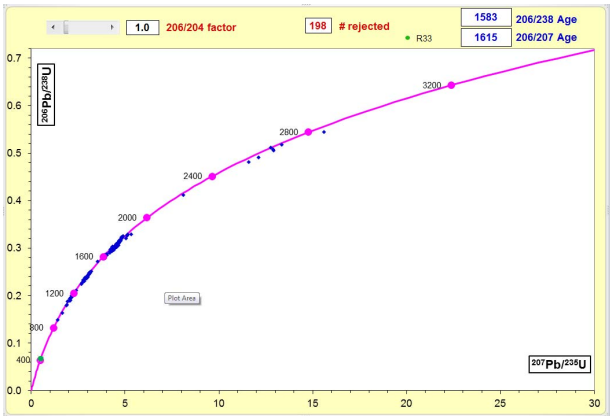


Figure 1f. *Concordia plot under a 5% discordance filter for sample BFD-105.*

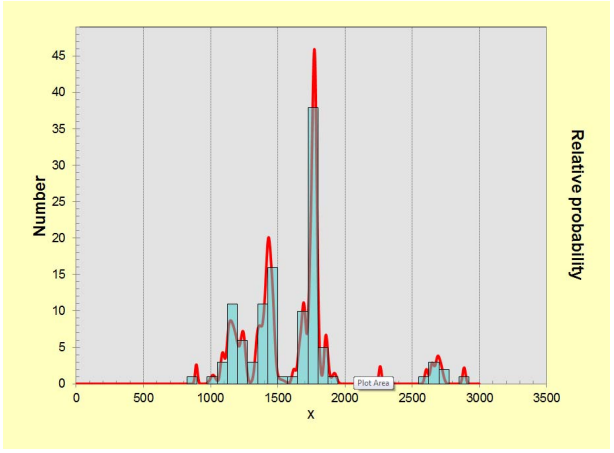


Figure 1d. *Histogram of sample BFD-105 under a 5% discordance filter.*

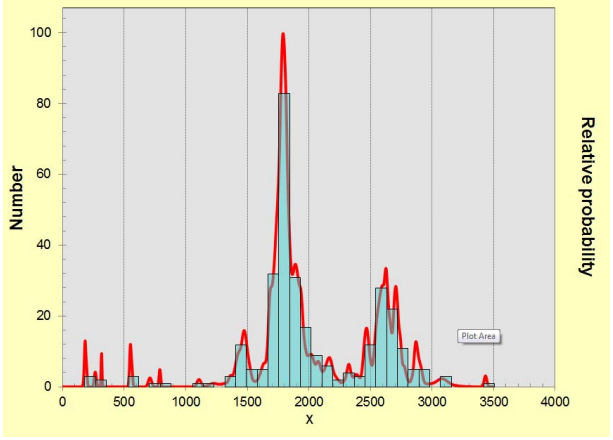


Figure 2a. *Histogram of sample BFD-107a under a 70% discordance filter.*

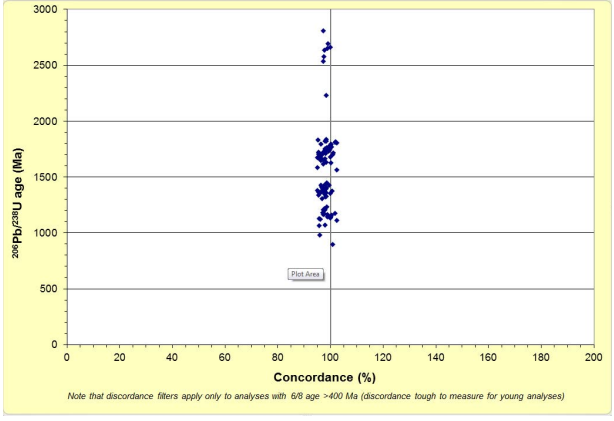


Figure 1e. *Plot of age vs. concordance for sample BFD-105 under a 5% discordance filter.*

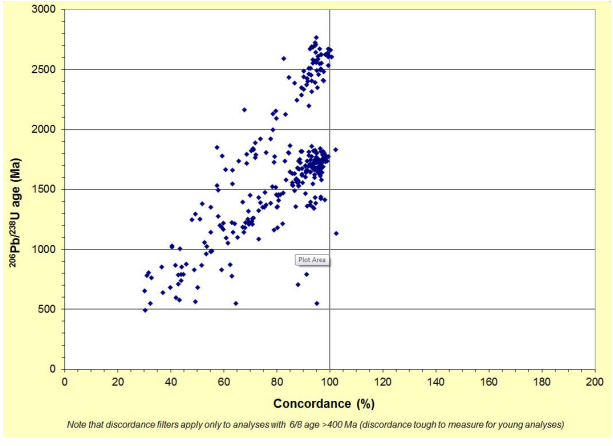


Figure 2b. *Plot of age vs. concordance for sample BFD-107a under a 70% discordance filter.*

around 2.2 Ga and a possible very discordant 3.2 Ga peak. Now at a 5% discordance filter, BFD-105 has a small 2.6-2.87 Ga peak, a big 1.77 Ga peak, a medium 1.43 Ga, and a small 1.14 Ga age peaks (Figure 1d, 1e, 1f). BFD-107a has about 2 grains at 3.43 Ga, a big 2.45-2.68 Ga peak range, a big 1.79 Ga peak, a small 1.48 Ga peak, and about 4 grains at 1.10 Ga (Figures 2d, 2e, 2f). The most significant difference is the 1.1 Ga age population, which is absent in sample BFD-107a, though there are subtle differences in the maximum peak ages and relative proportions of other age populations.

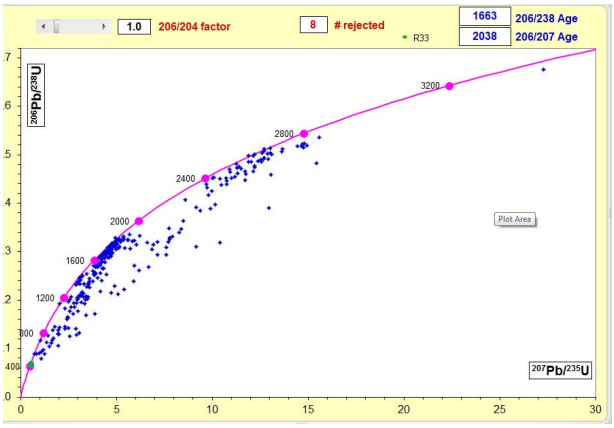


Figure 2c. *Concordia plot under a 70% discordance filter for sample BFD-107a.*

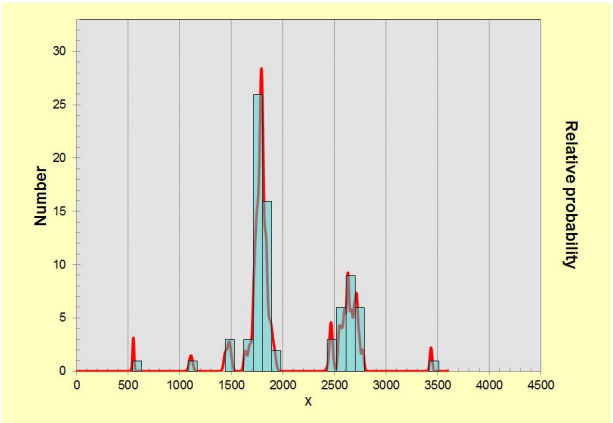


Figure 2d. *Histogram of sample BFD-107a under a 5% discordance filter. Here we notice the absence of the 1.14 Ga age peak present in BFD-105.*

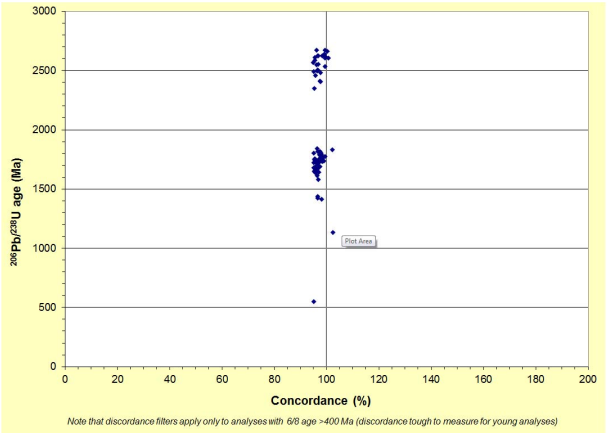


Figure 2e. *Plot of age vs. concordance for sample BFD-107a under a 5% discordance filter. Similarly in this plot, we notice the absence of the 1.14 Ga age peak in sample BFD-107a while there is no gap in sample in BFD-105.*

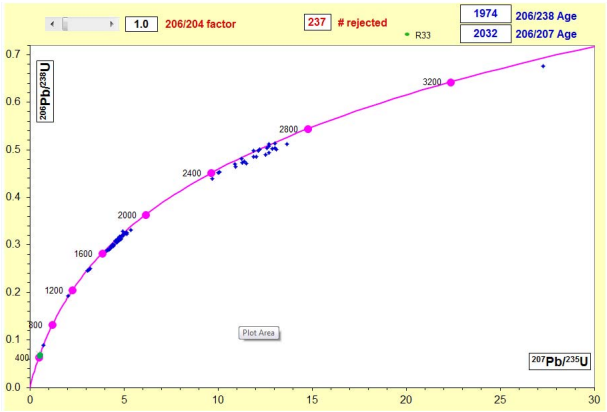


Figure 2f. *Concordia plot under a 5% discordance filter for sample BFD-107a.*

STATISTICAL ANALYSIS

Further interpretation was done using multidimensional scaling analysis which displays relations of objects by displaying distances between them connected by cords of varying patterns (Chen, 2016). In Figure 3, relations between several pendants are displayed. Samples BFD-105 and BFD-107a are marked as G1 and G2 respectively. Based on the solid connecting line and proximity of the samples, BFD-105 (G1) shows great similarity to R3, a sample of the Lower Vinini Formation (Linde et al., 2016). BFD-105 (G1) is also similar to C5, which is BFD-101 (a neighboring pendant). BFD-107a

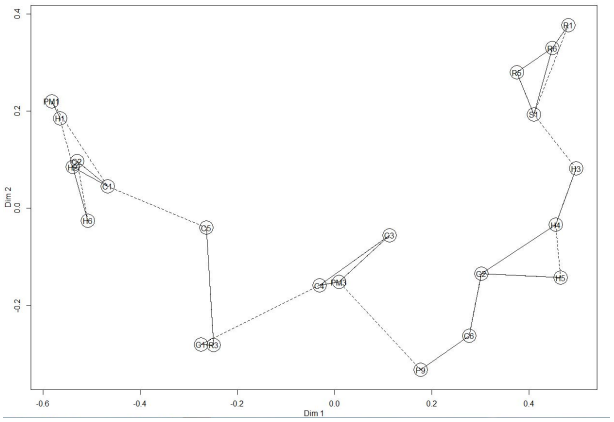


Figure 3. *Plot created using multidimensional scaling analysis to display similarities between several samples including BFD-105 (G1) and BFD-107a (G2).*

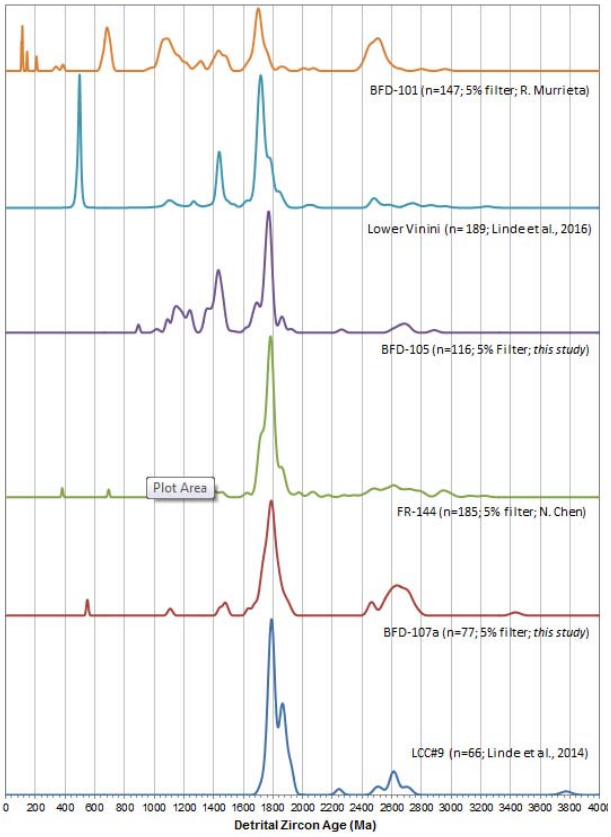


Figure 4. *Normalized probability density plot showing similarities between pendants in the peaks that they share. The top three pendants show similarities with each other while the bottom three also showcased similarities with each other. Both sets show differences in peaks ranging from 1.1 Ga to 1.5 Ga.*

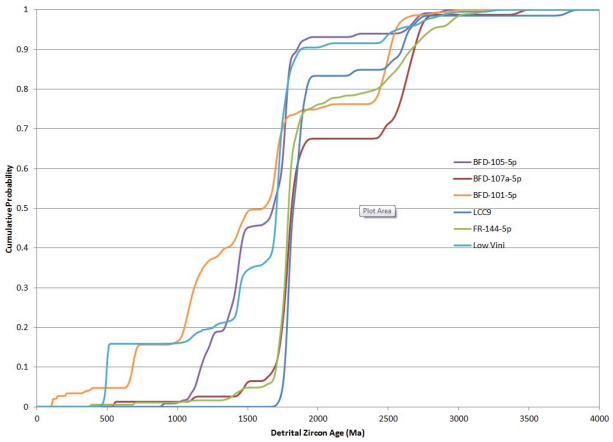


Figure 5. *Cumulative probability density plot showing the percentage of grains that are similar between age peaks of different samples.*

(G2) displays a high similarity to two samples from the Harmony B Formation (H4 (LCC#4) and H5(LCC#9); Linde, 2017) and to a pendant in Sequoia National Park (C6; Chen, 2016).

These comparisons are facilitated by plotting cumulative and normalized probability density plots. Figure 4 shows the age peaks for all samples that show similarities to BFD-105 and BFD-107a. This normalized probability density plot shows the relationships of the several pendants in the similarities of peaks that they share. Noticeably, the 1.1 peak is not present within BFD-107a, FR-144, and LCC#9 but is present within samples BFD-105, BFD-101, and the Lower Vinini Formation. The cumulative plot compares the percentage of grains that are actually present in each age peak, and compares different samples (Figure 5): e.g., curves for
BFD-107a, FR-144, and LCC#9 are sub-parallel, differing only slightly with regard to the percentatge of grains defining each age peak.

CONCLUSION

In the Sierra Nevada Mountain range, heat and stress associated with arc magmatism have greatly changed the pre-Mesozoic strata, rendering most dating techniques useless. However, with new U-Pb dating techniques on zircon grains from clastic sedimentary rocks, it

is possible to determine the maximum depositional age of the unmetamorphosed protoliths and describe their provenance. Such techniques were applied to the Fairview and Cowcreek pendants and lead to the determination that both pendants are different in terms of their maximum depositional age and provenance. One interpretation is that a cryptic fault juxtaposed the protoliths of the two pendants prior to their metamorphism by Mesozoic arc magmas.

REFERENCES

Chen, N., 2016, Detrital Zircon Uranium-Lead Geochronology of the Schoonover Sequence (Golconda Allochthon) and the Yokuts Pendants (Wester Sierra Nevada Mountains): Insight into the Late Paleozoic Evolution of Western Laurentia (M.S. Thesis): Fullerton, California State University, 96 p.

Lund, K., Aleinikoff, J.N., Evans K.V., duBray, E.A., Dewitt, E.H., Unruh, D.M., 2010, SHRIMP U-Pb dating of recurrent Cryogenian and Late Cambrian-Early Ordivician alkalic magmatism in central Idaho: Implication for Rodinian rift tectonics: Geological Society of America Bulletin, v. 122, p. 430-453.

Linde, G.M., Texler Jr., J.H., Cashman, P.H., Gehrels, G., Dickinson, W.R., 2017, Three- Demensional Evolution of the early Paleozoic Western Laurentia Margin: New insights from Detrital Zircon U-Pb Geochronology and HF Isotope Geochemistry of the HarmonyFormation of Nevada: Tectonics, v. 36, p. 2347 – 2369.

Linde, G.M., Trexler Jr., J.H., Cashman, P.H., Gehrels, G., Dickinson, W.R., 2016, Detrital Zircon U-Pb and Hf Isotope geochemistry of the Roberts Mountain allochthon: New Insight into the early Paleozoic tectonics of western North America: Geosphere, v. 12, p. 1016-1031.

Ross, D.C., 1987, Metamorphic Framework Rocks of the Southern Sierra Nevada, California: United States Department of the Interior Geologic Survey.

Conditions And Persistence Of Mcdonald Well, Bird Spring, And North Bird Spring In The Mojave Desert, San Bernardino County, California

Eddie Reyes

Advisors: [1]Jeffrey R. Knott, Ph.D., and [2]Anna L. Garcia

[1]*Department of Geological Sciences, California State University, Fullerton*

[2]*Mojave Water Agency, Apple Valley, CA*

ABSTRACT

Desert springs are often the sole sources of water and biodiversity across large areas of arid land. Their importance cannot be overlooked in conservation efforts and water resource management because they provide research opportunities for studying evolution and adaptation. We measured discharge, temperature, salinity, conductivity, and pH at three Mojave Desert springs: McDonald Well, Bird Spring, and North Bird Spring from March 2018 to March 2019. McDonald Well is a regional spring with discharge throughout. Lower discharge measurements following significant rain events signify a lag time between local precipitation and discharge. Higher salinity and conductivity indicate significant water-rock interaction and a long recharge path. Discharge gradually increasing with steady rainfall implies that discharge is directly related to local precipitation, but not reliant on it. Bird Spring is the least persistent of the three springs. Discharge as drips was observed only once. Salinity and conductivity measurements were fresher than McDonald Well, implying less rock-water interaction and a short recharge path. Thus, Bird Spring discharge is reliant entirely on local precipitation. Discharge from North Bird Spring was observed three times with discharge characterized as fracture seepage and dripping from the tuff cliff face. Low salinity and conductivity support low rock-water interaction and a short recharge path. Because Bird Spring and North Bird Spring do not see a substantial amount of discharge, restoring the springs to their natural state by the removal of man-made diversion would not harm the environment; biota cannot utilize the springs much in the spring's current state regardless. Removal of McDonald Well's diversion would not be harmful because its steady discharge ensures that the spring will reestablish itself.

Extending our view of a shallow-dipping Sierra Nevada Fault System south of Mount Whitney near Lone Pine, California

Jackson Flanagan
Advisor: Phillip A. Armstrong, Ph.D.
Department of Geological Sciences, California State University, Fullerton

The uplift of the Sierra Nevada Mountains has been debated for over 100 years, yet the orientation of the fault system that accommodates this uplift, the Sierra Nevada Frontal Fault System (SNFFS), is poorly constrained. The SNFFS, is the western boundary of Basin and Range Province normal faulting (Figure 1). Kinematic models of uplift, extension, and slip rates generally are made using an assumed 60° dipping SNFFS based on normal Andersonian fault mechanics. However, until recently, there have been no studies that directly measure the orientation of SNFFS faults to confirm or refute the assumption of a 60° dip.

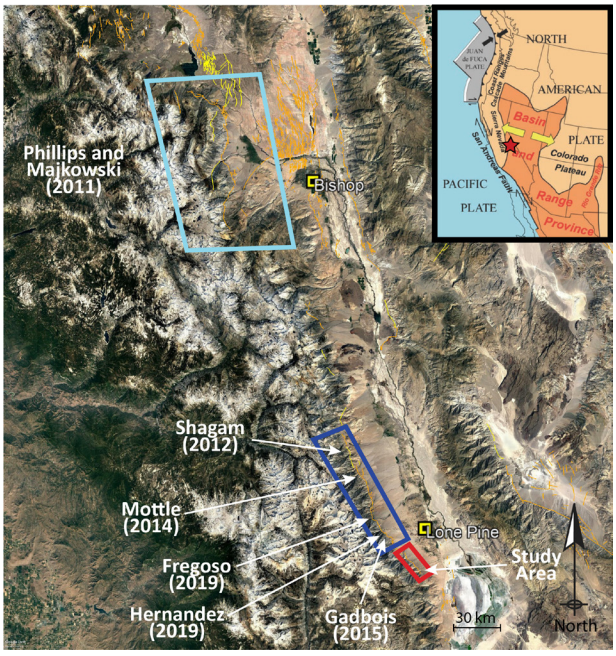


Figure 1. *Google Earth image of Owens Valley area. Holocene fault traces (yellow) and Late Quaternary fault traces (orange) from USGS Quaternary Fault and Fold database. Phillips and Majkowski (2011) field area shown in light blue. Field areas for Hernandez (2019), Fregoso (2019), Gadbois (2015), Mottle (2014), and Shagam (2012) outlined in dark blue. Inset map is adapted from Lillie (2005) and shows regional tectonics and extension of Basin and Range.*

An earlier analysis of the SNFFS by Phillips and Majkowski (2011) (Figure 1) shows that SNFFS faults near Bishop dip 25 - 35°. To further evaluate the orientation of the SNFFS farther south near Lone Pine (Figure 1), two faults between Carroll Creek and Diaz Creek were surveyed. Differential GPS was used to survey the fault exposures during low light conditions. The survey included 26 survey points along 1.5 km of fault exposure at Carroll Creek and 18 survey points along 2 km of exposure at Diaz Creek. Plane-fitting software was used to evaluate the x-y-z data from GPS surveys to produce a best-fit fault orientation. The Carroll Creek fault section strikes N24W and dips30°E (Figure 2), the section at Diaz Creek fault section strikes N41W and dips 22°E. Additionally, a direct exposure of the SNFFS (Figure 3) was located near Carroll Creek. The fault here offsets Quaternary alluvial fan deposits a minimum of 6.5 m. Direct dip measurements of the exposed fault produced a dip of 30°.

The results of this study are consistent with other SNFFS fault studies completed by CSUF geology students in the Lone Pine to Independence area (figure 1). These studies conclude: (1) the Tuttle Creek section strikes N34W and dips 36°E (Hernandez, 2019); (2) the Whitney Portal Road section strikes N24W and dips 28°E (Fregoso, 2019); (3) the Tuttle Creek bedrock section strikes N12W and dips 30°E (Gadbois, 2015); (4) the Bairs Creek section strikes N20W and dips 23°E (Mottle, 2014); (5) the Shepard Creek section strikes N44W and dips 34°E (Shagam, 2012); and (6) the Independence Creek section strikes N12W and dips

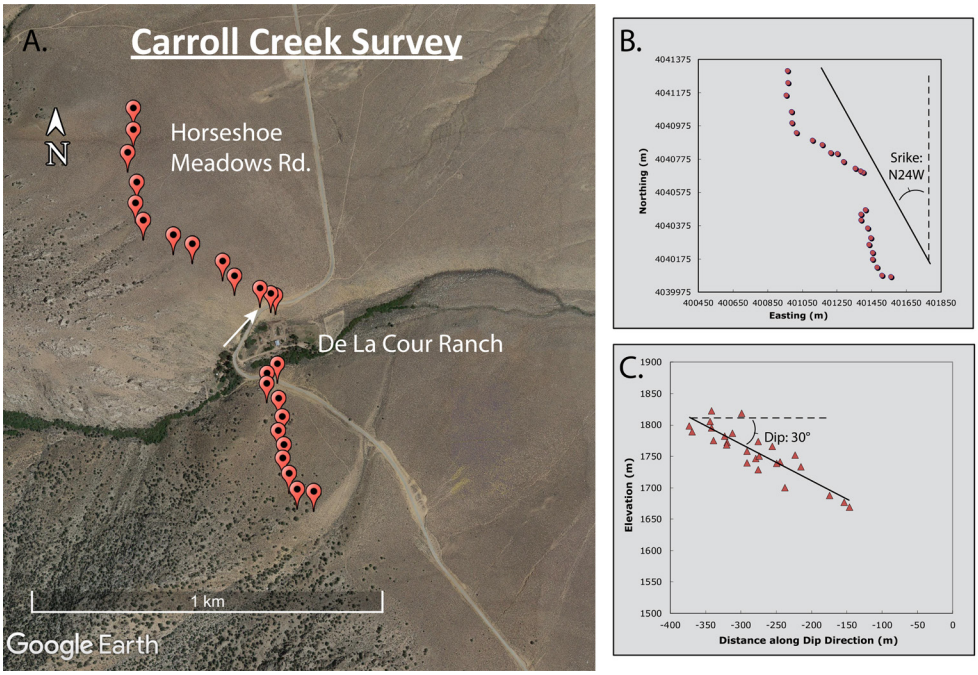


Figure 2. *GPS survey results for Carroll Creek field area. (A) Google Earth image with GPS survey points superimposed over field area. (B) UTM coordinates of mapped fault trace showing strike line that produces the best-fit dip using plane fitting routine. (C) Distance versus elevation data showing best-fit dip for survey results from plane fitting routine. White arrow in (A) shows location of the exposed fault surface in figure 3.*

Figure 3. *Field photo of Carroll Creek road cut fault exposure. Arrows point towards the fault trace. A marker unit offset by the fault is highlighted in purple. View is to the south. Direct measurement of this fault exposure produces dip of 30°E.*

29°E (Shagam, 2012). Thus, when combined with the results of Phillips and Majkowski (2011), the >100 km of the SNFFS from south of Lone Pine to north of Bishop appears to dip ~25 – 35°E.

Fault dip plays an important role in evaluating kinematic factors such as fault slip rate, extension and uplift rate and magnitudes, and comparisons of long- and short-term rates. Shallow dipping faults can significantly affect extension and slip rates derived from fault profile or total basin to range crest vertical offset estimates. Le et al. (2007) modeled the horizontal extension rates of the SNFFS near Lone Pine using a fault profile analysis and found Pleistocene horizontal extension rates of 0.2 mm/yr. These extension rates of Le et al. (2007) are increased by a factor of three to four when revised to account for 30° dipping faults.

REFERENCES

Fregoso, E., in preparation, Orientation of the Sierra Nevada Frountal Fault System in the vicinity of Whitney Portal Road, California. Unpublished undergraduate thesis.

Gadbois, B., 2015, Fault orientations of the Sierra Nevada Frontal Faulr Zone in the vicinity of Lone Pine, California: implications for basin and range extension. Unpublished undergraduate thesis.

Hernandez, J., in preparation, Low angle normal faults of the Sierra Nevadas near Lone Pine, California. Unpublished undergraduate thesis.

Lee, J., Le, K., Owen, L. A., and Finkel, R., 2007, Late Quaternary slip rates along the Sierra Nevada frontal fault zone, California: slip partitioning across the western margin of the Eastern California Shear Zone–Basin and Range Province: Geologic Society of America Bulletin, v. 119, no. 1-2, p. 240-256.

Lillie, R., J., 2005, Parks and Plates: The Geology of Our National Parks, Monuments, and Seashores, W. W. Norton & Company, 298 p.

Mottle, G., 2014, Evaluation of the Sierra Nevada Frontal Fault System at Bairs Creek in the vicinity of Manzanar, California. Unpublished undergraduate thesis.

Phillips, F. M., and Majkowski, L., 2011, The role of low-angle normal faulting in active tectonics of the northern Owens Valley, California: Lithosphere, v. 3, no. 1, p. 22-36.

Shagam G., 2012, Orientation of the Sierra Nevada Frontal Fault Zone near Independence and Lone Pine, California: Unpublished undergraduate thesis.

The Nature of Textural and Compositional Variability in the Lake Vernon Granodiorite in Western Yosemite National Park, CA.

Jamie Hayward and Cullen Scheland
Advisor: Valbone Memeti, Ph.D.
California State University Fullerton, Department of Geological Sciences

ABSTRACT

Granodioritic Sierran plutons are often characterized by within-unit, local to regional textural and compositional (modal) variability, too subtle for separation into mapping units. These within-unit changes might be due to differences in cooling histories, magma mixing, internal fractionation, and crystal accumulation. This study uses cathodoluminescence (CL) and petrography on plagioclase and its abundance in 7 samples from across the 97-95 Ma Lake Vernon Granodiorite (Klv), Jack Main Canyon intrusive suite, western Yosemite National Park, CA, to determine the nature of this variability.

Three plagioclase populations were identified. In order of abundance they are: (1) grains (150-3000 μm) that exhibit oscillatory and patchy zoning in both crossed-polarized light (XPL) and CL, 95% of grains have albitic overgrowth; (2) grains (150-2000 μm) that have no zoning (XPL and CL); and (3) grains (350-3000 μm) with oscillatory zoning characterized by thicker rims and inclusions of hornblende and opaque minerals (XPL). All samples display cumulate textures of populations 1 and 2 over 5-25% of the thin section area. K-feldspar crystallized interstitially adjacent to the plagioclase cumulates in 4 of 7 samples; the other 3 have K-feldspar grown in interstitial pockets throughout the thin section. The samples with the highest degree of plagioclase accumulation (15-25%) are located at the NE gradational contact with an older plutonic unit. The samples with the highest abundance of population 2 are found at the gradational contacts with an older plutonic unit (NE), and a younger plutonic unit (SW). Population 3 has no spatial preference.

Our results suggest that 3 petrographically distinct plagioclase populations likely originated in an earlier Klv magma or from adjacent units and concluded their magmatic histories together in the Klv. Cumulate textures suggest that melt migrated at least locally. More abundant plagioclase accumulation at Klv boundaries suggests boundary effects led to greater interstitial melt loss. These findings suggest Klv variability was caused by changing conditions throughout crystallization and/or magma mixing and accumulation. The presence of 3 populations in all samples across the Klv suggests that the Klv was melt interconnected across the pluton and underwent thorough crystal mixing.

Using Holocene sediments from North Yolla Bolly Lake in the Northern Coast Range (CA) to investigate the California precipitation dipole

Judith A. Avila, [1]Joseph A. Carlin, [2]Glen M. Macdonald, [1]Alexandre Woodward, [1]Jenifer Leidelmeijer, [2]Jiwoo Han, and [2]Benjamin Nauman

Advisor: [1]Matthew E. Kirby, Ph.D.

[1]*Department of Geological Sciences, California State University, Fullerton*

[2]*University of California, Los Angeles*

ABSTRACT

Where California receives its precipitation is as important, if not more, than how much precipitation it receives. Key to the "where and how much" issue is the position and strength of the California precipitation dipole (Dettinger et al., 1998; Wise, 2010). This dipole defines an approximate boundary of opposite precipitation signals – wetter north/drier south and vice versa. Based on a 500-yr tree ring study, the average position of the dipole sits at 40° N latitude (Wise, 2016). However, over 500 years, the dipole has varied between 35° and 44° N. To extend the dipole's record beyond 500 years, we are collecting lake sediment cores along a north-south coastal CA transect. Here, we present initial results from North Yolla Bolly Lake (40° N), located in the Northern Coast Range. Six sediment cores were collected in July 2018 from various water depths. Here, I focus on core NYBLRC18-6 from the modern littoral zone in 0.37 m of water. In general, the sediments consist of variegated brown-colored, organic-rich muds with occasional laminae. An age model was developed using Bacon v 2.2, based on six radiocarbon measurements on discrete organic materials and an assumed surface age of 2018 A.D. The total sediment history is approximately 10k cal yrs BP for core NYBLRC18-6. A variety of sediment analyses were determined including water content, dry bulk density, magnetic susceptibility, total organic matter, total carbonate, and grain size. Increases in percent sand is used to infer changes in the flux of terrigenous detritus during above average wet winters and/or discrete precipitation events. Changes in percent total organic matter and percent clay are used to infer long-term variations in lake level and lake trophic state. Results will be compared to existing sites north and south of North Yolla Bolly Lake to produce a behavioral reconstruction of the California precipitation dipole for the Holocene.

Clast Provenance And Detrital Zircon Dating Of Pliocene Fluvial Conglomerate: Implications For Paleohydrology Of Crooked Creek And Deep Springs Valley, California

Lindsey M. Langer

Advisors: [1]Jeffrey R. Knott, Ph.D., [1]Diane Clemens-Knott, Ph.D., and [2]Anna Garcia

[1]*Department of Geological Sciences, California State University, Fullerton,*

[2]*Mojave Water Agency, Apple Valley, CA*

ABSTRACT

The paleohydrology of the western Great Basin has long been of interest to geologists and biologists interested in reconstructing dispersal pathways for various biota. Past studies briefly describe a fluvial conglomerate, with an interbedded 3.1 Ma tuff, that crops out on the 270-m-high ridge separating Deep Springs Valley from Eureka Valley. Additional inspection reveals that the fluvial conglomerate contains boulder-to gravel-sized clasts in a sandy matrix interbedded with cross-bedded coarse sandstone. Dominant clast types are sandstone, limestone, quartz monzonite, granite, and basalt. U-Th/Pb dates of 273 detrital zircons separated from the sandstone matrix are dominantly Jurassic (183.6–167.6 Ma), with an average zircon age of 175.33±0.38 Ma that correlates with local plutons. Lesser Triassic and Cretaceous zircon are also present. Six Proterozoic grains are likely derived from Paleozoic metasedimentary sources as they yield U-Th/Pb ages with known North American provenance: ca. 1.1, 1.4, and 1.8 Ga. Zircon ages and clast types can be used to identify the likely drainage basin from which the conglomerate was sourced. Basalt is not found in Wyman Creek drainage, eliminating that drainage. The drainage network of modern Crooked Creek, however, is underlain by Jurassic granodiorite, basalt and Paleozoic sedimentary rocks making it the likely source of the conglomerate. I infer that (a) the conglomerate provenance is to the north-northwest in the White Mountains across Deep Springs Valley, and that (b) post-3.1 Ma uplift along the Deep Springs fault cut off the northwest-to-southeast river flowing in modern Crooked Creek and formed Deep Springs Valley.

Incorporating Migration into an Age-Structured Population Model with Variable Mortality Rate

Gabriel Martinez and Freddy Nungaray
Advisors: Laura Smith Chowdhury, Ph.D.
Department of Mathematics, California State University, Fullerton

ABSTRACT

We modify the McKendrick Von-Foerster population model, an age-structured population model, by including a migration term to model population dynamics when migration is a significant contributor to population size. Migration is treated as a continuous function of age and time. We consider the model's mortality rate as both a constant and as a function of age. By allowing the mortality rate to be a function of age, the model better captures more realistic death rates. We then apply this model to the United States population. When the mortality rate is considered as a constant, the model has 4.85% relative error. When the mortality rate is considered as a function of age, it yields a better approximation with a 2.07% relative error.

1 INTRODUCTION

Immigration and emigration have often become an important topic for countries with large numbers of migrants. In 2015, it was estimated that 244 million people were migrating into a foreign country seeking employment, refuge, and other incentives [8]. Over three million Venezuelans have emigrated to seek economic opportunity [6]. The civil war in Syria encouraged people to seek refuge in countries along the Mediterranean Sea [6]. Among all the countries in the world, the United States has become the country with the most immigrants [8]. The United States is highly sought by asylum seekers, evidenced recently by a large migrant caravan of roughly 7,000 central Americans who wish to settle in the country [9]. In order to have an accurate population count for a given region, population models are often employed. However, many population models focus on closed systems where individuals do not migrate in or out of the system. Since migration can greatly affect the size of a population, it will be necessary to include it in population models for countries where migration is significant, such as the United States.

Mathematical modeling of populations is used to analyze changes in a population. Typically, a population model includes an initial population, birth rates, death rates, and other factors relevant to changes in the population. We modified a population model, known as the McKendrick Von-Foerster age-structured population model, to include migration [19]. An age-structured population model assumes individuals of the same age have the same growth rate and death rate. Thus, these models use age and time as independent variables, often resulting in a partial differential equation (PDE) model. An age-structured population model is more appropriate for consideration since there is significant variability in the age of migrants. Therefore, we need to consider age as an independent variable for our model.

Another consideration for using an age-structured population model is the mortality rate for humans, which is highly dependent on age. For example, among U.S. children aged 1–4 years there were 28.6 deaths per 100,000 in 2007 [16], which is an extremely low death rate compared to older age groups. Health problems such as heart disease and cancer remain as the top two causes of deaths for adults [7], giving a significantly higher death rate for the elderly. This leads us to include a mortality rate that depends on age in our age-structured population model. With this age dependent mortality function and migration, we are able to produce reasonable population estimates when applied to the United States population.

2 LITERATURE REVIEW

Population models are often represented by differential or difference equations. One simple model is the Malthusian growth model, which assumes a population's growth rates and death rates are proportional to the population's current size, leading to possibly unconstrained exponential growth [10]. The logistic growth model incorporates a carrying capacity for which the environment can sustain [10]. In addition, rival or predatory species may affect the population. The Lotka-Volterra model, often referred to as the predator-prey model, investigates the interaction between a particular prey and predator population [17].

Another class of models are compartment models, often used for epidemiology. An SIR model, standing for susceptible, infected, and recovered individuals, depicts how populations change between classes for a given disease [20]. There are several different types of these compartment models, which depend on the states of the disease and the modeling assumptions. Often, simplifying assumptions are made to make a mathematical model that is easier to solve or analyze. These models often assume an independent variable of time only, resulting in systems of ordinary differential equations (ODEs).

More complex models that allow for additional considerations often result in PDE models. This allows for more factors to make the models more realistic, however, they become more difficult to solve. Some population models include independent variables such as age, size, or location. For example, the McKendrick model investigates the change in cell population relative to age and size [19]. James W. Sinko and William Streifer proposed a size-structured population model that uses time, size, and age as independent variables [11]. The McKendrick Von-Foerster model uses time and age as independent variables [19] and is a special case of the Sinko-Streifer model.

Population models are often represented by discrete models. For example, the Leslie model is used to track the succession of population groups over a selected interval of time [13]. The Leslie model can also numerically approximate the McKendrick Von-Foerster equation [13]. Sinko and Striefer analyzed the Leslie model in the study of Daphnia pulex and found it to be inadequate to accurately depict the behavior of smaller animals [13].

3 U.S. POPULATION DATA

Data were collected for the following different categories: population, natality, mortality, and migration. We include these categories in order to address the elements presented in the proposed model in Section 4.2.

3.1 Population Data

The population data were retrieved from the national intercensal datasets 2000–2010 and the annual estimates of the resident population 2010–2016 [1, 2]. The national intercensal datasets are intercensal estimates of resident population organized by single age per year, sex, race, and Hispanic origin from April 1, 2000 to July 1, 2010. The estimates are calculated using a population base and the result of the population size from the year 2000 from the previous decennial census, and then updated using the Das Gupta method, which assumes that the ratio from the intercensal estimate to the postcensal estimate follows a geometric progression [3]. The population data includes births and deaths of U.S. residents, domestic migration, international migration, and net movement of the U.S. armed forces [3].

The annual estimates of the resident population for the datasets from April 1, 2010 to July 1, 2016 consisted of two categories: age and sex of the individual. These estimates take into account the change in population from births, deaths, and migration. The annual population estimates are calculated using a cohort component method derived from a demographic balancing equation using a population base from the previous decennial census estimate [4]. These estimates add net migration, which include international and domestic, but assumes the sum to be zero, so it does not affect population change [4].

The national intercensal datasets and the annual estimates datasets are combined and organized by age for each year from 2004 to 2014. However, since these are two separate datasets with two different methods of estimating population size, the intercensal data provide population estimates up to age 85, and the remaining estimates are summed into age

999, an unrealistic age. The annual estimates provide data by age for each year up to age 100 with individuals over 100 grouped together. The population estimate from the intercensal dataset for age 85 ranges from 4,545,883 to 5,367,301 individuals from 2004 to 2009, while the annual estimates mark the population size for age 85 to fall between 915,071 to 914,112 individuals. This presents an unexplained jump for this particular age. Due to this inconsistency in the data, this study utilizes only the data from ages 0 to 84 from the combined datasets.

3.2 Birth Data

The birth data were obtained for the years 2004 to 2014 from the Centers for Disease Control and Prevention WONDER online data bank [5]. The datasets, natality for 2003–2006 and the natality for 2007–2016, were queried to extract the birth totals grouped by year. The data account for births in the United States from U.S. residents and non-residents [5].

3.3 Mortality Data

The mortality data were obtained from the CDC WONDER online databases for underlying causes of death from 1999 to 2016. Totals were grouped by ages starting with "$<$ 1 year", "1–4 years", and then grouped by 5 year intervals. The data do not include deaths of nonresidents. The death counts are derived from death certificates from all fifty states, excluding Puerto Rico, Guam, the Virgin Islands, and other U.S. territories [5]. Death counts that are zero to nine are suppressed.

3.4 Migration Data

Immigration data were retrieved from the U.S. Homeland Security Yearbook of Immigration Statistics [14]. The data group by year contains data on foreign nationals who were granted lawful permanent residence, admitted with a temporary status, or applied for asylum or refugee status [14]. Persons admitted with temporary status are considered "nonimmigrants" since their visit is short and without the intention of permanently residing in the U.S. territories. For the purposes of this study, legal permanent residents, refugees, and asylees are considered in the estimates of immigration data, as represented. The data are grouped by 5-year age groups up to age group 65-74. Furthermore, to include the nonimmigrant population would require data on their departure, which is not recorded by any U.S. agency. Since their stay is considered temporary, primarily less than one year, their inclusion in the population count is not significant. We consider the immigration data to be estimates since the U.S. Homeland Security Yearbooks acknowledge that revisions occur to eliminate duplicates [14].

4 METHODOLOGY

This section describes our methodology for incorporating migration into an age-structured population model. We first introduce size and age-structured population models and then provide the proposed model.

4.1 The Sinko-Streifer and the McKendrick Von-Foerster Population Model

The Sinko-Streifer equation is a size-structured population model given as a partial differential equation with time and size as independent variables [10]. The Sinko-Streifer equation assumes individuals of the same size share some common traits [10]. The model defines $u(t, x)$ as the number of individuals of size x at time t. Individuals who increase in size can transition into other sizes as time progresses. It is also assumed that individuals of the same size have the same growth rate and the same probability of death. The birth rate is assumed to be proportional to the size of the population [10]. Also, the size of the population is assumed to be significantly large, and the size x of an individual is

$$
\begin{aligned}
\gamma &= 137464.1856, \\
\beta &= 0.0019, \\
\sigma &= 34.7412, \\
\lambda &= 0.0083.
\end{aligned}
\tag{4}
$$

The immigration data and the migration model, $M(t, \alpha)$, are shown in Figure 1. A superimposed comparison is shown in Figure 2. The relative error of the data and model is 0.0827.

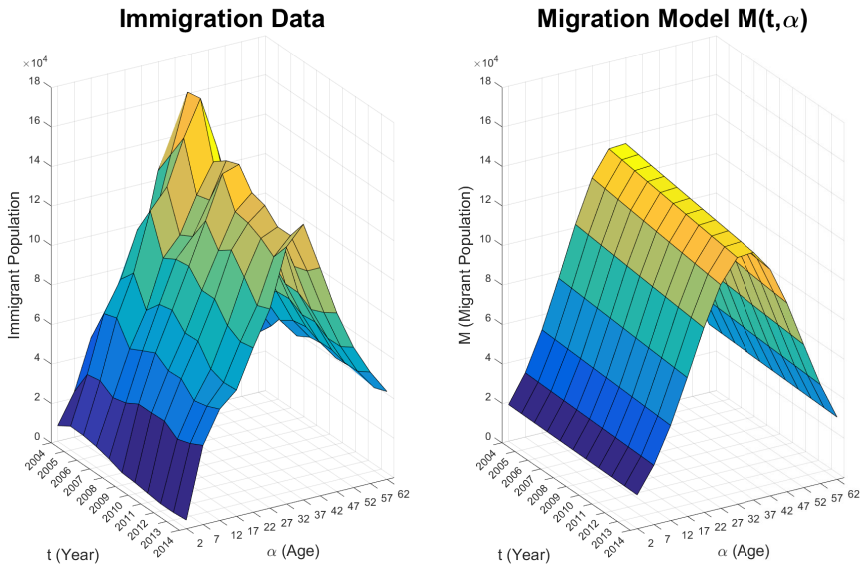


Figure 1: *The surface graph on the left represents the data for the estimated immigrant population gathered from the U.S. Homeland Security Immigrant Yearbook Statistics. The surface graph on the right was created using Equation 3.*

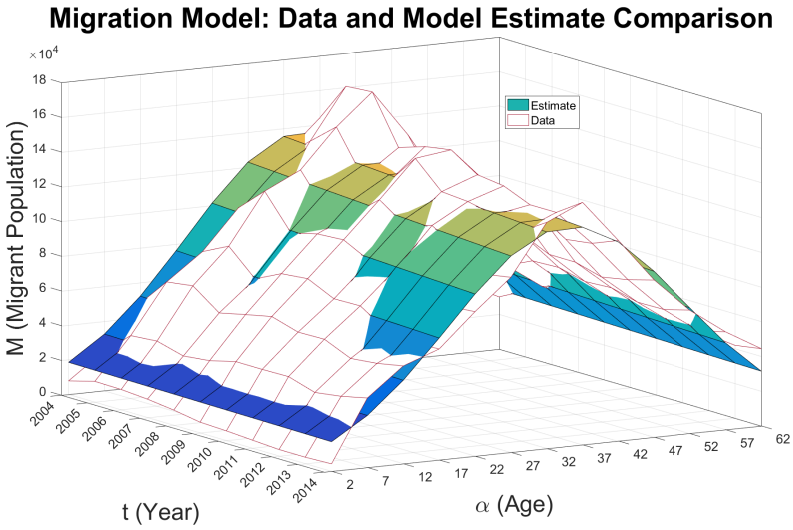


Figure 2: *The migration data and the migration model $M(t, \alpha)$ are overlaid for a better comparison.*

5 IMPLEMENTATION

This section gives the model implementation details. To find an approximate solution, we apply the method of characteristics [10]. We consider two different cases regarding $\mu(t, \alpha)$, the mortality rate. We initially consider the mortality rate to be a constant. Then, we look at the case when the mortality rate is a function of age. The solution will be divided into two parts, the initial condition driven solution and the recruitment driven solution [10]. The initial condition driven solution comes from the population presented initially at t_0. The recruitment driven solution comes from the population born into it during the time period of interest.

5.1 Applying the Method of Characteristics

Using the method of characteristics, we take our PDE and convert it into three ordinary differential equations,

$$\begin{aligned}
\frac{dt}{ds} &= 1, \text{ and thus } t = s + t_0, \\
\frac{d\alpha}{ds} &= 1, \text{ and thus } \alpha = s + \alpha_0, \\
\frac{du}{ds} &= -\mu u + M.
\end{aligned}$$

Solving for s giving the results,

$$s = t - t_0 \text{ and } \alpha = (t - t_0) + \alpha_0.$$

Next for $\frac{du}{ds}$, along with the explicit function $M = M(t, \alpha)$ from Equation 3, we get

$$\frac{du}{ds} + \mu u = \gamma e^{-\beta(\alpha-\sigma)^2} e^{\lambda s}. \tag{5}$$

The solution to this ODE depends on the form of $M(t, \alpha)$ which we will consider by case.

5.2 Constant Mortality Rate $\mu(t,\alpha) = \mu_0$

We first consider $\mu(t,\alpha)$ as a constant rate. For $\mu(t,\alpha) = \mu_0$, Equation 5 becomes

$$\frac{du}{ds} + \mu_0 u = \gamma e^{-\beta(s+\alpha_0-\sigma)^2+\lambda s}.$$

We first find the initial condition driven solution. The solution to the ODE is

$$e^{\mu_0 s}u = \gamma e^{k_1}\int e^{-\beta s^2+k_2 s}ds,$$

where $k_1 = -\beta(-2\alpha_0\sigma+\alpha_0^2+\sigma^2)$ and $k_2 = -2\beta\alpha_0+2\beta\sigma+\lambda+\mu_0$ with parameters γ, β, σ, and λ given by Equation 4. In solving for μ_0, the integral $\int e^{-\beta s^2+k_2 s}ds$ cannot be solved analytically. We use a Taylor Series expansion to approximate the integral. Then the population u is given by

$$u = \gamma e^{k_1-\mu_0 s}\left[s + k_2\frac{s^2}{2} + (-2\beta+k_2^2)\frac{s^3}{6} + (-6\beta k_2+k_2^2)\frac{s^4}{24} + (12\beta(\beta-k_2^2)+k_2^4)\frac{s^5}{120}\right] + c_1\gamma e^{k_1-\mu_0 s}.$$

We solve for c_1 by considering the initial condition for $u_0 = u(0,\alpha)$, the initial population at time t_0. This yields the initial condition driven solution for the constant mortality function,

$$u(t,\alpha) = \gamma e^{k_1-\mu_0 s}\left[s + k_2\frac{s^2}{2} + (-2\beta+k_2^2)\frac{s^3}{6} + (-6\beta k_2+k_2^3)\frac{s^4}{24} + (12\beta(\beta-k_2^2)+k_2^4)\frac{s^5}{120}\right] + u_0 e^{-\mu_0 s}.$$

Similarly, we find the recruitment driven solution. Let $R(t)$ be the total number of individuals born into the population at time t. Then the recruitment driven solution is given by

$$u(t,a) = \gamma e^{k_1-u_0 s}\left[s + k_2\frac{s^2}{2} + (-2\beta+k_2^2)\frac{s^3}{6} + (-6\beta k_2+k_2^3)\frac{s^4}{24} + (12\beta(\beta-k_2^2)+k_2^4)\frac{s^5}{120}\right] + R(t)e^{-\mu_0 s}.$$

5.3 Mortality Rate as a Function of Age, $\mu(t,\alpha) = \mu(\alpha)$

The mortality data from the CDC WONDER databases are given in five year age groups along with the total population for each age group per year [3, 4]. Using these two data categories, we obtain an approximate rate of the number of deaths per population size for each age group per year, visually represented in Figure 3. We fit a polynomial to these rates to approximate the $\mu(\alpha)$ function. Having fewer parameters will help us implement the model more feasibly, so the approximation was given by the polynomial of sixth degree,

$$\mu(\alpha) = l_6\alpha^6 + l_5\alpha^5 + l_4\alpha^4 + l_3\alpha^3 + l_2\alpha^2 + l_1\alpha + l_0 \tag{6}$$

with the approximated coefficients shown in Table 1.

Equation 6 estimates the mortality rate according to age. We apply an integrating factor to Equation 5 with $\mu = \mu(\alpha)$ to get

$$\frac{d}{ds}\left(e^L u\right) = \gamma e^{\lambda s-\beta(s+\alpha_0-\sigma)^2+L},$$

where $L(s) = \frac{l_6}{7}(s+\alpha_0)^7 + \frac{l_5}{6}(s+\alpha_0)^6 + \frac{l_4}{5}(s+\alpha_0)^5 + \frac{l_3}{4}(s+\alpha_0)^4 + \frac{l_2}{3}(s+\alpha_0)^3 + \frac{l_1}{2}(s+\alpha_0)^2 + l_0 s$, obtained by integrating $\mu(\alpha)$. This results in the initial condition driven solution,

Table 1: Approximated and Optimized Parameters for $\mu(\alpha)$

	Approximate	**Optimized**
l_0	3.26792 $\times 10^{-3}$	4.6660 $\times 10^{-5}$
l_1	-1.06115 $\times 10^{-3}$	-1.8441 $\times 10^{-3}$
l_2	1.18189 $\times 10^{-4}$	1.6218 $\times 10^{-4}$
l_3	-5.68820 $\times 10^{-6}$	-6.1658 $\times 10^{-6}$
l_4	1.34161 $\times 10^{-7}$	1.2327 $\times 10^{-7}$
l_5	-1.51054 $\times 10^{-9}$	-1.2714 $\times 10^{-9}$
l_6	6.609 $\times 10^{-12}$	5.482 $\times 10^{-12}$

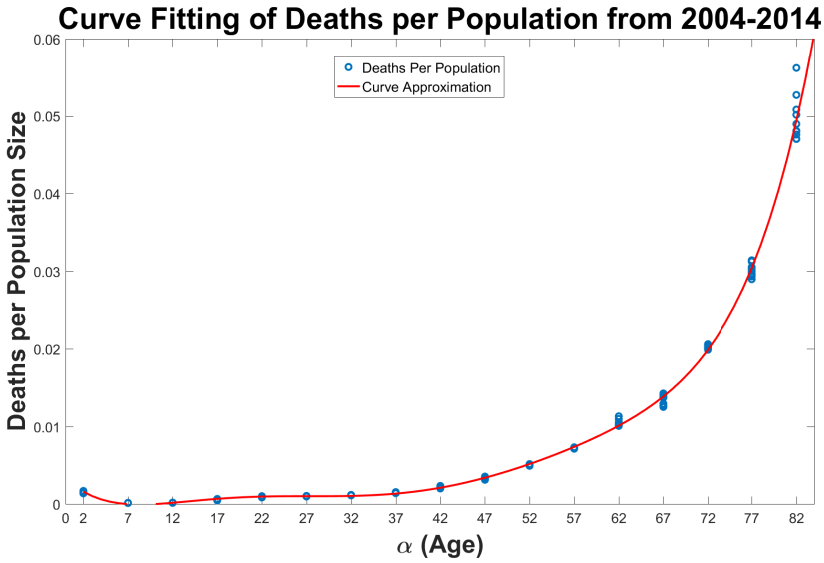


Figure 3: *The approximate death rate for the median age of each age group was calculated for each year. A sixth degree polynomial that represents the mortality rate function for the U.S. population from 2004--2014 is plotted with the data. The polynomial describes the mortality rate as a function of age.*

and the recruitment driven solution

$$u(t,\alpha) = \gamma e^{-L(t)} \int^t e^{\lambda s - \beta(s+\alpha_0-\sigma)^2 + L(t)} ds + e^{L(0)} u(0,\alpha) e^{-L(t)} \tag{8}$$

in Equation 3, giving the McKendrick Von-Foerster equation, the optimal $\mu_0 = -0.00014$. The mortality rate value produced is unrealistic to the context of this study, so it must be rejected. An unrealistic mortality rate indicates the need for migration to be included in the model.

Next, using $\mu(\alpha)$ from Equation 6 with the approximation parameters of Table 1 gives the estimate for $u(t, \alpha)$ shown in Figure 5. We find that the model approximates the population data with a relative error of 0.0274. To see if a better polynomial would result in a closer approximation, we use the least-squares error minimization method to optimize the coefficients of $\mu(\alpha)$, giving the optimized coefficients provided in Table 1. The population estimate with this optimized polynomial has a 0.0102 relative error, and is shown in Figure 6. The optimal polynomial is plotted in Figure 7. However, the curve yields $\mu(\alpha) < 0$ for certain values of α which is unrealistic.

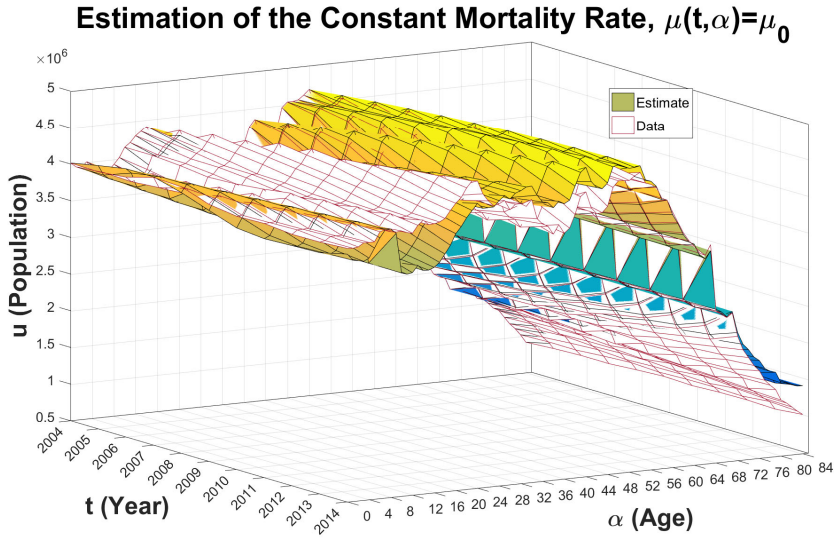


***Figure 4**: Population estimate for a constant mortality rate. The constant mortality rate is estimated to be $\mu_0 = 0.0195315$. The model has a relative error of 0.0485. The estimate and data are super imposed in this figure as a visual reference.*

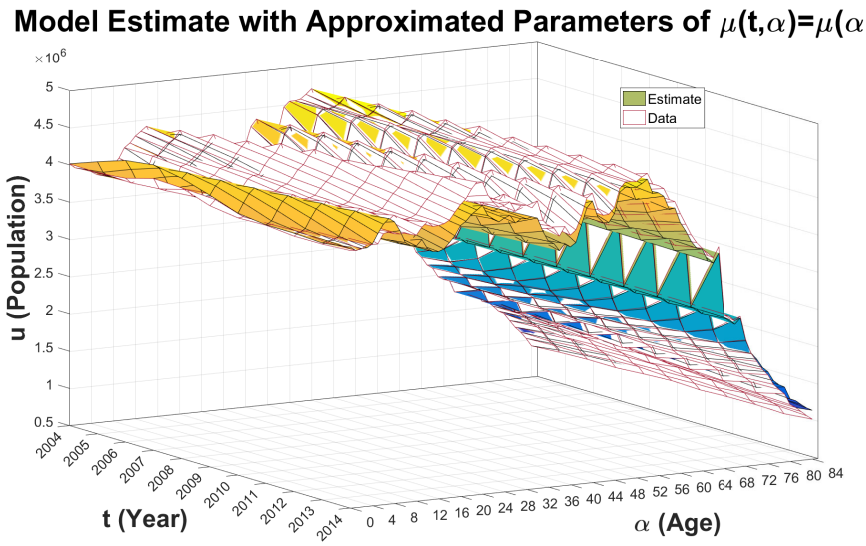


Figure 5: *Population estimate using the sixth degree polynomial $\mu(\alpha)$ with the approximated parameter values shown in Table 1. This estimate produces a relative error of 0.0274 to the overall U.S. population.*

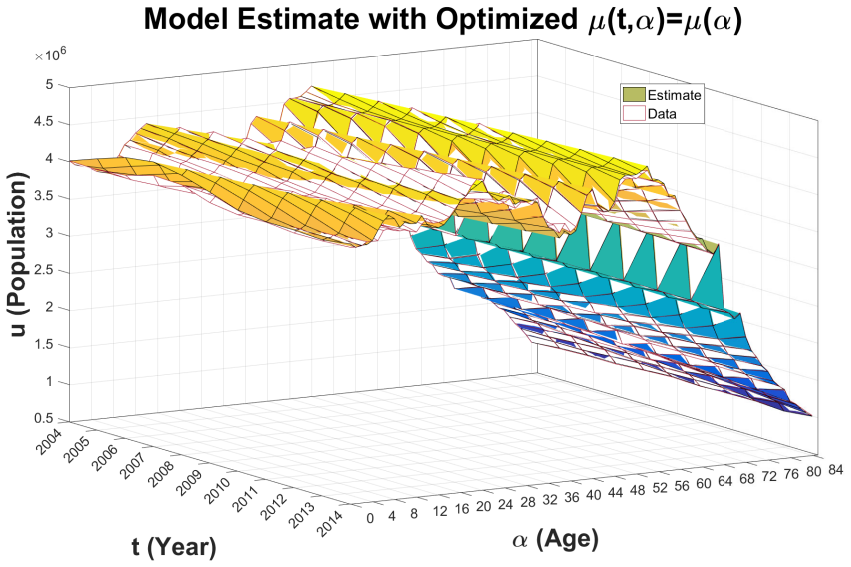


Figure 6: *The mortality rate $\mu(t,\alpha) = \mu(\alpha)$ is approximated by a sixth degree polynomial, which is used to estimate $u(t,\alpha)$. The coefficients of $\mu(\alpha)$ are optimized by a least-squares error minimization method to produce the estimate of the population with a relative error of 0.0102.*

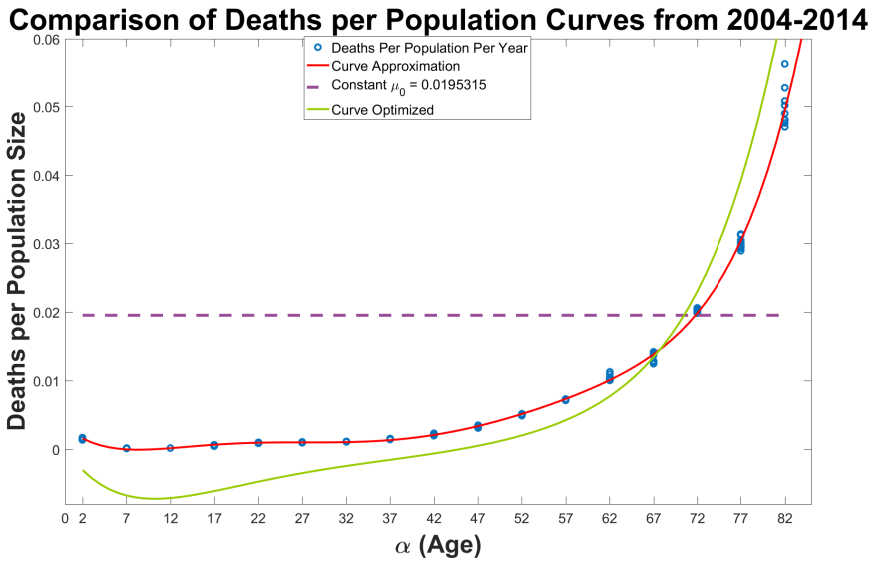


Figure 7: *A comparison of the approximated parameter values for the polynomial $\mu(\alpha)$ (red curve) and the optimized values (green curve). The constant mortality rate is shown relative to the varying mortality rate. The optimized coefficients for $\mu(\alpha)$ produced a more accurate estimation for the U.S. population with a relative error of 0.0102. However, the optimized polynomial curve offsets the data of deaths per population with negative values for people less than 67 years old.*

7 ANALYSIS

Our proposed model was able to capture the population dynamics of an open system. When we considered mortality

REFERENCES

[1] Population Division U.S. Census Bureau, available online at the URL: https://www.census.gov/data/ datasets/timeseries/ demo/ popest/ intercensal-2000-2010-national.html

[2] Population Division U.S. Bureau, available online at the URL: https://factfinder.census.gov/faces/tableservices/jsf/pages/ productview.xhtml?pid= PEP_2016_PEPSY ASEXNprodT ype = table

[3] Population Division U.S. Census Bureau, available online at the URL: https://www2.census.gov/programssurveys/popest/ technical-documentation/methodology/intercensal/2000-2010-intercensal-estimatesmethodology.pdf

[4] Population Division U.S. Census Bureau, available online at the URL: https://www2.census.gov/programssurveys/popest/ technical-documentation/methodology/2010-2016/2016-natstcopr-meth.pdf

[5] Center for Disease Control and Prevention, available online at the URL: http://wonder.cdc.go/ucd-icd10.html

[6] Central Intelligence Agency World Factbook available online at the URL: https://www.cia.gov/library/publications/the-world-factbook/geos/ve.html

[7] J. Xu, M.D. Sherry, L. Murphy, B.S. Kenneth, D. Kochanek, M.A. Brigham, B.S. Bastian, E. Arias, Deaths: Final data for 2016, *Natl Vital Stat Rep* , **67** (2018), 1–76.

[8] US News Washington D.C., available online at the URL: https://www.usnews.com/news/slideshows/10-countriesthat-take-the-most-immigrants?slide=11

[9] BBC News, available at the online URL: https://www.bbc.com/news/world-latin-america-45951782

[10] H.T. Banks, H.T. Tran, *Mathematical and Experimental Modeling of Physical and Biological Processes*, Chapman & Hall/CRC, 2009.

[11] J.W. Sinko, S. William, A New Model for Age-Size Structure of a Population, *Ecol.* **48** (1967) 910–918.

[12] B.L. Keyfitz, N., Keyfitz, The McKendrick Partial Differential Equation and the Uses in Epidemiology and Population Study, *Math and Computational Modeling* **26** (1997) 1--9.

[13] P.H. Leslie, On the use of matrices in certain population mathematics, *Biometrika*, **33** (1945), 183--212.

[14] Department of Homeland Security, avialable online at the URL: https://www.dhs.gov/immigrationstatistics/ yearbook

[15] C. Akselrud, A.E. Punt, C.F. Lee, The affects of applying mis-specified age- and size-structured population models, *Fisheries Research*, **188** (2017), 58--73.

[16] G.K. Singh, P.C. Van Dyck, Infant mortality in the United States, 1935--2007: Over seven decades of progress and disparities, *U.S. Dept. HHS*, (2010), 1--8.

[17] Predetor-Prey Population Dynamics: the Lotka-Volterra model in Stan, available online at the URL: https://mc-stan.org/users/documentation/case-studies/lotka-volterra-predator-prey.html

[18] C.J.R. Hogue, C. Vasques, Toward a Strategic Approach for Reducing Disparities in Infant Mortality, *Am J Public Health*, **92** (2002), 552-–556.

[19] A.G, McKendrick. Applications of mathematics to medical problems, *PEMS*, **44** (1925), 94 -- 130.

[20] M. Martcheva, An Introduction to Mathematical Epidemiology. *Introduction to Epidemic Modeling*, Springer 2015.

[21] S.J. Farlow, *Partial Differential Equations for Scientists and Engineers*, John Wiley & Sons, 1982.

[22] L Samuel, L.N Borrell, Body mass index categories and mortality risk in US adults: the effect of overweight and obesity on advancing death, *AJPH*, **104** (2014), 512 - 519.

Invertible Polynomials and their Mirrors

Kenneth Eaves
Advisor: Christopher Lyons, Ph.D.
Department of Mathematics, California State University, Fullerton

INTRODUCTION

Our research focuses on certain polynomials called invertible polynomials and how they are related to one another through a mirror operation. The interest in this mirror operation comes most notably from the attempts of physcists to describe a certain phenomenon known as mirror symmetry [1]. Each invertible polynomial is quasihomogeneous and thus has a weight system associated with it; conversely each of these weight systems can have numerous associated invertible polynomials.

Our research began by investigating how the mirror operation relates these weight systems to one another. Drawing from computer investigations we develop and prove 2 theorems in this direction. It turns out that there is a certain connection between n-variable invertible polynomials and $(n-1)$-variable invertible polynomials that can explain some of the relations we found. Using a process called specialization, we start with an $(n-1)$-variable invertible polynomial and relate it to multiple n-variable invertible polynomials coming from different weight systems.

PRELIMINARIES

We begin with some examples of invertible polynomials:

- Chain type: $x_1^{a_1}x_2 + \cdots + x_{n-1}^{a_{n-1}}x_n + x_n^{a_n}$ where $n \geq 1$.
- Loop type: $x_1^{a_1}x_2 + \cdots + x_{n-1}^{a_{n-1}}x_n + x_n^{a_n}x_1$ where $n \geq 1$

The special case of chain type with $n = 1$ is also called a Fermat Type.
While the definition of invertible polynomials is somewhat involved, the following theorem [2] says they're built up from the examples above.

Theorem 2.1 (Kreuzer-Skarke). Any invertible polynomial is a direct sum of polynomials that are each of chain type or loop type.

By direct sum we mean the variables in one atomic type do not appear in another atomic type and these atomic types are added together.

Definition 2.2. Consider the polynomial

$$\begin{aligned} F(x_1, x_2, \cdots, x_n) &= x_1^{a_{1,1}}x_2^{a_{1,2}}\cdots x_n^{a_{1,n}} + x_1^{a_{2,1}}x_2^{a_{2,2}}\cdots x_n^{a_{2,n}} + \cdots + x_1^{a_{m,1}}x_2^{a_{m,2}}\cdots x_n^{a_{m,n}} \\ &= \sum_{i=1}^{m} x_1^{a_{i,1}}\cdots x_n^{a_{i,n}} \\ &= \sum_{i=1}^{m}\prod_{j=1}^{n} x_j^{a_{i,j}}. \end{aligned}$$

which has n variables and m monomials. The exponent matrix of F is

$$A(F) = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}.$$

Definition 2.3. $F(x_1, x_2, \cdots, x_n)$ is quasihomogeneous with weight system $\{w_1, w_2, \cdots, w_n : d\}$ if

$$F(\lambda^{w_1} x_1, \lambda^{w_2} x_2, \cdots, \lambda^{w_n} x_n) = \lambda^d F(x_1, x_2, \cdots, x_n).$$

Definition 2.4. Let F and $\tilde{F}$ be invertible polynomials. If the exponent matrix of F equals the transpose of the exponent matrix of $\tilde{F}$, we call F and $\tilde{F}$ mirror polynomials.

Example 2.5. Consider

$$F(x_1, x_2, x_3, x_4) = x_1^2 x_3 + x_2^3 x_4 + x_3^5 + x_4^7 x_2$$

and notice that

$$F(\lambda^4 x_1, \lambda^3 x_2, \lambda^2 x_3, \lambda x_4) = \lambda^{10} F(x_1, x_2, x_3, x_4)$$

for any constant λ. F is quasihomogeneous with weight system $\{4, 3, 2, 1 : 10\}$. We note F has summands of both chain type and loop type. Let $A(F)$ be the exponent matrix of F, then

$$A(F) = \begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & 3 & 0 & 1 \\ 0 & 0 & 5 & 0 \\ 0 & 1 & 0 & 7 \end{bmatrix}.$$

Now the mirror polynomial $\tilde{F}$ of F has exponent matrix

$$A(\tilde{F}) = A(F)^T = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 1 \\ 1 & 0 & 5 & 0 \\ 0 & 1 & 0 & 7 \end{bmatrix},$$

and thus

$$\tilde{F}(x_1, x_2, x_3, x_4) = x_1^2 + x_2^3 x_4 + x_1 x_3^5 + x_2 x_4^7.$$

MOTIVATING EXAMPLE

Consider the weight system $\{21, 14, 6, 1 : 42\}$. One can show there are four invertible polynomials with this weight system.

$$\begin{aligned} F_1 &= x_1^2 + x_2^3 + x_3^7 + x_4^{42} \\ F_2 &= x_1^2 + x_2^3 + x_3^7 + x_4^{36} x_3 \\ F_3 &= x_1^2 + x_2^3 + x_3^7 + x_4^{28} x_2 \\ F_4 &= x_1^2 + x_2^3 + x_3^7 + x_4^{21} x_1 \end{aligned}$$

Let $\tilde{F}_i$ denote the mirror polynomial of F_i for $1 \leq i \leq 4$. Then

$$\begin{aligned}
\tilde{F}_1 &= x_1^2 + x_2^3 + x_3^7 + x_4^{42} \\
\tilde{F}_2 &= x_1^2 + x_2^3 + x_3^7 x_4 + x_4^{36} \\
\tilde{F}_3 &= x_1^2 + x_2^3 x_4 + x_3^7 + x_4^{28} \\
\tilde{F}_4 &= x_1^2 x_4 + x_2^3 + x_3^7 + x_4^{21}.
\end{aligned}$$

Notice that $F_1 = \tilde{F}_1$ and so naturally they would both have the same weight system, but this is not the case for the other three polynomials. Their weight systems are as follows:

$$\begin{aligned}
\tilde{F}_1 &: \{21, 14, 6, 1 : 42\} \\
\tilde{F}_2 &: \{18, 12, 5, 1 : 36\} \\
\tilde{F}_3 &: \{14, 9, 4, 1 : 28\} \\
\tilde{F}_4 &: \{10, 7, 3, 1 : 21\}.
\end{aligned}$$

It just so happens that each of these three weight systems have 4 invertible polynomials associated with them,

Table 1. *All invertible polynomials associated with the specified weight system.*

{18,12,5,1:36}	**{14,9,4,1:28}**	**{10,7,3,1:21}**
$G_1 = x_1^2 + x_2^3 + x_3^7 x_4 + x_4^{36}$	$H_1 = x_1^2 + x_2^3 x_4 + x_3^7 + x_4^{28}$	$J_1 = x_1^2 x_4 + x_2^3 + x_3^7 + x_4^{21}$
$G_2 = x_1^2 + x_2^3 + x_3^7 x_4 + x_4^{31} x_3$	$H_2 = x_1^2 + x_2^3 x_4 + x_3^7 + x_4^{24} x_3$	$J_2 = x_1^2 x_4 + x_2^3 + x_3^7 + x_4^{18} x_3$
$G_3 = x_1^2 + x_2^3 + x_3^7 x_4 + x_4^{24} x_2$	$H_3 = x_1^2 + x_2^3 x_4 + x_3^7 + x_4^{19} x_2$	$J_3 = x_1^2 x_4 + x_2^3 + x_3^7 + x_4^{14} x_2$
$G_4 = x_1^2 + x_2^3 + x_3^7 x_4 + x_4^{18} x_1$	$H_4 = x_1^2 + x_2^3 x_4 + x_3^7 + x_4^{14} x_1$	$J_4 = x_1^2 x_4 + x_2^3 + x_3^7 + x_4^{11} x_1$

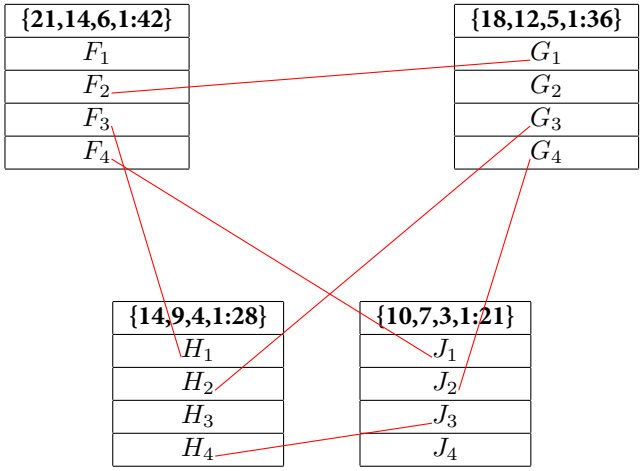
As we defined it, $\tilde{F}_2 = G_1$, $\tilde{F}_3 = H_1$, and $\tilde{F}_4 = J_1$, but what's interesting and not a priori obvious in this example, and what has motivated the rest of this paper, is the fact that the collection of 16 polynomials above is closed under the mirror operation. This is detailed in the table 2:

Table 2. *The polynomials mentioned are closed under the mirror operation.*

Weight System	**Polynomial**	**Mirror**	**Weight System**
$\{21, 14, 6, 1 : 42\}$	F_1	F_1	$\{21, 14, 6, 1 : 42\}$
$\{21, 14, 6, 1 : 42\}$	F_2	G_1	$\{18, 12, 5, 1 : 36\}$
$\{21, 14, 6, 1 : 42\}$	F_3	H_1	$\{14, 9, 4, 1 : 28\}$
$\{21, 14, 6, 1 : 42\}$	F_4	J_1	$\{10, 7, 3, 1 : 21\}$
$\{18, 12, 5, 1 : 36\}$	G_2	G_2	$\{18, 12, 5, 1 : 36\}$
$\{18, 12, 5, 1 : 36\}$	G_3	H_2	$\{14, 9, 4, 1 : 28\}$
$\{18, 12, 5, 1 : 36\}$	G_4	J_2	$\{10, 7, 3, 1 : 21\}$
$\{14, 9, 4, 1 : 28\}$	H_3	H_3	$\{14, 9, 4, 1 : 28\}$
$\{14, 9, 4, 1 : 28\}$	H_4	J_4	$\{10, 7, 3, 1 : 21\}$
$\{10, 7, 3, 1 : 21\}$	J_4	J_4	$\{10, 7, 3, 1 : 21\}$

Table 3 shows the connections a bit better. A line between the two polynomials indicate that the two are mirror polynomials.

Table 3. *A visual representation of the 16 polynomials and their associated mirrors.*



Note that F_1, G_2, H_3 and J_4 are their own mirrors.

SPECIALIZATIONS AND DECORATIONS

While observing the 16 polynomials mentioned in the motivating example, one may notice how similar each of the polynomials are. To capture these impressions precisely, we define below the specialization and decoration of a polynomial. Using a specific specialization, we can relate an n-variable invertible polynomial to an $(n-1)$-variable invertible polynomial. In the next section we will see that by a specific specialization of all of the polynomials in the motivating example, we get various decorations of the same 3-variable polynomial.

Definition 4.1. Let $F = F(x_1, \ldots, x_n)$. The i^{th} specialization of F is the polynomial $F_{(i)}$ defined by

$$\begin{aligned} F_{(i)}(x_1, \cdots, \hat{x_i}, \cdots, x_n) &= F(x_1, \cdots, 1, \cdots, x_n) \\ &= \lim_{x_i \to 1} F(x_1, \cdots, x_n). \end{aligned}$$

Definition 4.2. Let $f = f(x_1, \ldots, x_m)$ and $f^{(j)}$. For $0 \leq j \leq m$, the j^{th} decoration of f is:

- $f^{(j)} = f + x_j, \quad 1 \leq j \leq m$
- $f^{(j)} = f + 1, \quad j = 0$

We note that if $A(F)$ is the exponent matrix of F then the exponent matrix $A(F_{(i)})$ is the same, but with the i^{th} column removed. Also if $A(f)$ is the exponent matrix of f then the exponent matrix $A(f^{(j)})$ is the same, but with an additional row of either all zeros, or all zeros and a single 1.

MOTIVATING EXAMPLE (CONTINUED)

When we take the specialization of $(F_i)_{(4)}$ for each $1 \leq i \leq 4$ we see that we get some decoration of $h(x_1, x_2, x_3) = x_1^2 + x_2^3 + x_3^7$:

Table 4. *The fourth specialization of each polynomial and it's resulting decoration. A line indicates the mirror polynomial.*

{21,14,6,1:42}
$(F_1)_{(4)} = x_1^2 + x_2^3 + x_3^7 + 1 = h^{(0)}$
$(F_2)_{(4)} = x_1^2 + x_2^3 + x_3^7 + x_3 = h^{(3)}$
$(F_3)_{(4)} = x_1^2 + x_2^3 + x_3^7 + x_2 = h^{(2)}$
$(F_4)_{(4)} = x_1^2 + x_2^3 + x_3^7 + x_1 = h^{(1)}$

{18,12,5,1:36}
$(G_1)_{(4)} = x_1^2 + x_2^3 + x_3^7 + 1 = h^{(0)}$
$(G_2)_{(4)} = x_1^2 + x_2^3 + x_3^7 + x_3 = h^{(3)}$
$(G_3)_{(4)} = x_1^2 + x_2^3 + x_3^7 + x_2 = h^{(2)}$
$(G_4)_{(4)} = x_1^2 + x_2^3 + x_3^7 + x_1 = h^{(1)}$

{14,9,4,1:28}
$(H_1)_{(4)} = x_1^2 + x_2^3 + x_3^7 + 1 = h^{(0)}$
$(H_2)_{(4)} = x_1^2 + x_2^3 + x_3^7 + x_3 = h^{(3)}$
$(H_3)_{(4)} = x_1^2 + x_2^3 + x_3^7 + x_2 = h^{(2)}$
$(H_4)_{(4)} = x_1^2 + x_2^3 + x_3^7 + x_1 = h^{(1)}$

{10,7,3,1:21}
$(J_1)_{(4)} = x_1^2 + x_2^3 + x_3^7 + 1 = h^{(0)}$
$(J_2)_{(4)} = x_1^2 + x_2^3 + x_3^7 + x_3 = h^{(3)}$
$(J_3)_{(4)} = x_1^2 + x_2^3 + x_3^7 + x_2 = h^{(2)}$
$(J_4)_{(4)} = x_1^2 + x_2^3 + x_3^7 + x_1 = h^{(1)}$

Take note of how, for a given weight system, the 4^{th} specializations of the mirror polynomials in that weight system are all the same decoration of the polynomial $h(x_1, x_2, x_3)$. This is detailed in the following table:

Table 5. *The fourth specializations of the mirror polynomials in the same weight system are all the same decoration of h.*

Weight System	**Polynomial**	**4^{th} Specialization**	**4^{th} Specialization of Mirror**
$\{21, 14, 6, 1 : 42\}$	F_1	$h^{(0)}$	$h^{(0)}$
$\{21, 14, 6, 1 : 42\}$	F_2	$h^{(3)}$	$h^{(0)}$
$\{21, 14, 6, 1 : 42\}$	F_3	$h^{(2)}$	$h^{(0)}$
$\{21, 14, 6, 1 : 42\}$	F_4	$h^{(1)}$	$h^{(0)}$
$\{18, 12, 5, 1 : 36\}$	G_1	$h^{(0)}$	$h^{(3)}$
$\{18, 12, 5, 1 : 36\}$	G_2	$h^{(3)}$	$h^{(3)}$
$\{18, 12, 5, 1 : 36\}$	G_3	$h^{(2)}$	$h^{(3)}$
$\{18, 12, 5, 1 : 36\}$	G_4	$h^{(1)}$	$h^{(3)}$
$\{14, 9, 4, 1 : 28\}$	H_1	$h^{(0)}$	$h^{(2)}$
$\{14, 9, 4, 1 : 28\}$	H_2	$h^{(3)}$	$h^{(2)}$
$\{14, 9, 4, 1 : 28\}$	H_3	$h^{(2)}$	$h^{(2)}$
$\{14, 9, 4, 1 : 28\}$	H_4	$h^{(1)}$	$h^{(2)}$
$\{10, 7, 3, 1 : 21\}$	J_1	$h^{(0)}$	$h^{(1)}$
$\{10, 7, 3, 1 : 21\}$	J_2	$h^{(3)}$	$h^{(1)}$
$\{10, 7, 3, 1 : 21\}$	J_3	$h^{(2)}$	$h^{(1)}$
$\{10, 7, 3, 1 : 21\}$	J_4	$h^{(1)}$	$h^{(1)}$

THEOREMS

The following two theorems are motivated by the main example discussed above. The first theorem explains that the i^{th} specialization of the mirror of F is the mirror of the i^{th} specialization of F. The second theorem explains that if the i^{th} specialization of two n-variable polynomials that share a weight system result in decorations of the same $(n-1)$-variable polynomial, then the i^{th} specialization of their mirror polynomials will be equal.

Theorem 6.1. Let $F(x_1, x_2, \ldots, x_n)$ be an invertible polynomial with mirror polynomial $\tilde{F}$ and suppose the specialization $F_{(i)}$ is a decoration of the polynomial $h = h(x_1, \cdots, \hat{x_i}, \cdots, x_n)$. Then $\tilde{F}_{(i)}$ is a decoration of $\tilde{h}$.

Proof. Let $F(x_1, x_2, \ldots, x_n)$ be an invertible polynomial. Without loss of generality, assume F is a direct sum of the atomic types in Theorem 2.1. We can express the exponent matrix of F as a block matrix where each block is comprised of a single atomic type. Each of these blocks has the form

$$\begin{bmatrix} a_{1,1} & 1 & 0 & \ldots & 0 & 0 \\ 0 & a_{2,2} & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & 0 & \ldots & a_{k-1,k-1} & 1 \\ a_{k,1} & 0 & 0 & \ldots & 0 & a_{k,k} \end{bmatrix}$$

where $a_{p,p} > 1$ and $a_{k,1} \in \{0, 1\}$. Since F is a sum of atomic types, after rearranging the summands we can assume that $F(x_1, x_2, \ldots, x_n) = G(x_1, \ldots, x_k) + J(x_{k+1} \ldots, x_n)$ where G is of atomix type and J is invertible and $1 \leq i \leq k$. Then

$$A(F) = \begin{bmatrix} A(G) & 0 \\ 0 & A(J) \end{bmatrix},$$

and $A(F_{(i)})$ is $A(F)$ with the ith column removed, while $A(G_{(i)})$ is $A(G)$ with the ith column removed. Thus

$$A(F_{(i)}) = \begin{bmatrix} A(G_{(i)}) & 0 \\ 0 & A(J) \end{bmatrix}$$

where

$$A(G_{(i)}) = \begin{bmatrix} a_{1,1} & 1 & \ldots & 0 & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & a_{i-1,i-1} & 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & a_{i,i+1} & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & a_{i+1,i+1} & 1 & \ldots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{k,1} & 0 & \ldots & 0 & 0 & 0 & 0 & a_{k,k} \end{bmatrix}.$$

By the hypothesis, $F_{(i)}$ is a decoration of h. Note that in $A(G_{(i)})$, and hence $A(F_{(i)})$, the only row to give $+1$ or $+x_j$, as required by definition of a decoration, is row i. Therefore removing row i from $A(F_{(i)})$ produces $A(h)$, the exponent matrix of h. Thus

$$A(h) = \begin{bmatrix} A(g) & 0 \\ 0 & A(J) \end{bmatrix}$$

where

$$A(g) = \begin{bmatrix} a_{1,1} & 1 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & a_{i-1,i-1} & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & a_{i+1,i+1} & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & & \vdots \\ a_{k,1} & 0 & \dots & 0 & 0 & \dots & 0 & a_{k,k}. \end{bmatrix}$$

The exponent matrix of $\tilde{F}$ is

$$\begin{aligned} A(\tilde{F}) &= A(F)^T \\ &= \begin{bmatrix} A(G)^T & 0 \\ 0 & A(J)^T \end{bmatrix} \\ &= \begin{bmatrix} A(\tilde{G}) & 0 \\ 0 & A(\tilde{J}) \end{bmatrix} \end{aligned}$$

where

$$A(\tilde{G}) = \begin{bmatrix} a_{1,1} & 0 & \dots & a_{k,1} \\ a_{1,2} & a_{2,2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & a_{k,k} \end{bmatrix}$$

Thus we get

$$A(\tilde{F}_{(i)}) = \begin{bmatrix} A(\tilde{G}_{(i)}) & 0 \\ 0 & A(\tilde{J}), \end{bmatrix}$$

where

$$A(\tilde{G}_{(i)}) = \begin{bmatrix} a_{1,1} & 0 & \dots & 0 & 0 & \dots & a_{k,1} \\ 1 & a_{2,2} & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & a_{i-1,i-1} & 0 & \dots & 0 \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & a_{i+1,i+1} & \dots & 0 \\ 0 & 0 & \dots & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & a_{k,k}. \end{bmatrix}$$

The i^{th} row of $A(\tilde{F}_{(i)})$ is again the only one to give $+1$ or $+x_j$ and so removing it gives us the exponent matrix of the unique polynomial of which $\tilde{F}_{(i)}$ is a decoration. Note that removing the i^{th} row of $A(\tilde{G}_{(i)})$ gives the matrix

$$\begin{bmatrix} a_{1,1} & 0 & \dots & 0 & 0 & \dots & a_{k,1} \\ a_{1,2} & a_{2,2} & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & a_{i-1,i-1} & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & a_{i+1,i+1} & \dots & 0 \\ 0 & 0 & \dots & 0 & a_{i+1,i+2} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & a_{k,k}. \end{bmatrix} \tag{6.1}$$

Now we see that (4.1) is equal to $A(g)^T = A(\tilde{g})$. So removing the ith row from $A(\tilde{F}_{(i)})$ gives

$$\begin{bmatrix} A(\tilde{g}) & 0 \\ 0 & A(\tilde{J}). \end{bmatrix}$$

Therefore $\tilde{F}_{(i)}$ is a decoration of $\tilde{h} = \tilde{g} + \tilde{J}$. □

Theorem 6.2. Suppose $F(x_1, x_2, \dots, x_n)$ and $G(x_1, x_2, \dots, x_n)$ are invertible polynomials with the same weights and degree. Suppose that for some i, $F_{(i)}$ and $G_{(i)}$ are decorations of $h(x_1, \dots, \hat{x}_i, \dots, x_n)$. Then $\tilde{F}_{(i)} = \tilde{G}_{(i)}$, i.e. $\tilde{F}_{(i)}$ and $\tilde{G}_{(i)}$ are the same decoration of $\tilde{h}$.

Proof. Consider two invertible polynomials F and G that both correspond to the weight system $\{w_1, w_2, \cdots, w_n : d\}$, and let $A(F)$ and $A(G)$ be the exponent matrices of F and G respectively. Through a permutation of the variables (columns of the exponent matrix) we can arrange for the last variable (last column) to be the specialization variable. Also we can rearrange terms of F so that it looks like

$$F = \cdots + x_n^{a_{n,n}} x_s^{e_s}$$

where x_n is the variable we are specializing, and $e_s \in \{0, 1\}$. Likewise

$$G = \cdots + x_n^{b_{n,n}} x_t^{e_t}$$

where x_n is the variable we are specializing, and $e_t \in \{0, 1\}$. The exponent matrices are

$$A(F) = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n-1} & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n-1} & a_{2,n} \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n-1} & a_{n,n} \end{bmatrix}$$

and

$$A(G) = \begin{bmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,n-1} & b_{1,n} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,n-1} b a_{2,n} & \\ \vdots & \vdots & & \vdots & \vdots \\ b_{n-1,1} & b_{n-1,2} & \cdots & b_{n-1,n-1} & b_{n-1,n} \\ b_{n,1} & b_{n,2} & \cdots & b_{n,n-1} & b_{n,n} \end{bmatrix}.$$

Suppose that $F_{(n)}$ and $G_{(n)}$ are both decorations of some polynomial $h(x_1, \ldots, x_{n-1}, \hat{x_n})$. Thus the (n, n)-minor of $A(F)$ is the exponent matrix of h, and likewise with $A(G)$. Thus,

$$A(h) = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n-1} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n-1} \\ \vdots & \vdots & & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n-1} \end{bmatrix}$$

$$= \begin{bmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,n-1} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,n-1} \\ \vdots & \vdots & & \vdots \\ b_{n-1,1} & b_{n-1,2} & \cdots & b_{n-1,n-1} \end{bmatrix}.$$

Thus,

$$a_{i,j} = b_{i,j} \text{ for all } 1 \leq i \leq n-1 \text{ and } 1 \leq j \leq n-1.$$

To show $\tilde{F}_{(n)} = \tilde{G}_{(n)}$ we must show $A(\tilde{F}_{(n)}) = A(\tilde{G}_{(n)})$, i.e. that the n^{th} column of $A(\tilde{F})$ is the same as the n^{th} column of $A(\tilde{G})$, less $a_{n,n}$ and $b_{n,n}$. We can reexpress the quasihomogenity of F in Definition 2.3 in terms of matrix multiplication.

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n-1} & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n-1} & a_{2,n} \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n-1} & a_{n,n} \end{bmatrix} \times \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_{n-1} \\ w_n \end{bmatrix} = \begin{bmatrix} d \\ d \\ \vdots \\ d \\ d \end{bmatrix}.$$

Multiplying the left side and setting each product equal to d gives us,

$$\begin{aligned} a_{1,1}w_1 + a_{1,2}w_2 + \cdots + a_{1,n-1}w_{n-1} + a_{1,n}w_n &= d \\ a_{2,1}w_1 + a_{2,2}w_2 + \cdots + a_{2,n-1}w_{n-1} + a_{2,n}w_n &= d \\ &\vdots \\ a_{n-1,1}w_1 + a_{n-1,2}w_2 + \cdots + a_{n-1,n-1}w_{n-1} + a_{n-1,n}w_n &= d \\ a_{n,1}w_1 + a_{n,2}w_2 + \cdots + a_{n,n-1}w_{n-1} + a_{n,n}w_n &= d \end{aligned}$$

We're not concerned about $a_{n,n}$ and so we can disregard the last equation, but for the others we solve for a_i, n for

$1 \leq n \leq n-1$:

$$\begin{aligned}
a_{1,n} &= \frac{d-(a_{1,1}w_1 + a_{1,2}w_2 + \cdots + a_{1,n-1}w_{n-1})}{w_n} \\
a_{2,n} &= \frac{d-(a_{2,1}w_1 + a_{2,2}w_2 + \cdots + a_{2,n-1}w_{n-1})}{w_n} \\
&\vdots \\
a_{n-1,n} &= \frac{d-(a_{n-1,1}w_1 + a_{n-1,2}w_2 + \cdots + a_{n-1,n-1}w_{n-1})}{w_n}.
\end{aligned}$$

Similarly we can show using $A(G)$:

$$\begin{aligned}
b_{1,n} &= \frac{d-(b_{1,1}w_1 + b_{1,2}w_2 + \cdots + b_{1,n-1}w_{n-1})}{w_n} \\
b_{2,n} &= \frac{d-(b_{2,1}w_1 + b_{2,2}w_2 + \cdots + b_{2,n-1}w_{n-1})}{w_n} \\
&\vdots \\
b_{n-1,n} &= \frac{d-(b_{n-1,1}w_1 + b_{n-1,2}w_2 + \cdots + b_{n-1,n-1}w_{n-1}+)}{w_n}.
\end{aligned}$$

Since $b_{i,j} = a_{i,j}$ for all $1 \leq i, j \leq n-1$, then we get

$$\begin{aligned}
a_{1,n} &= b_{1,n} \\
a_{2,n} &= b_{2,n} \\
&\vdots \\
a_{n-1,n} &= b_{n-1,n}.
\end{aligned}$$

Thus $A(\tilde{F}_{(i)} = A(\tilde{G}_{(i)}$, implying $\tilde{F}_{(i)} = \tilde{G}_{(i)}$ as required. □

ACKNOWLEDGEMENTS

I would like to thank the 2018 CSUF Math Summer Research Program for giving me the opportunity to further this research. The program also gave me the chance to present my research in a poster session, allowing me to entertain questions from interested students, faculty, and guest which helped to further my understanding and face unanswered questions.

REFERENCES

[1] P. Berglund and T. Hübsch [A generalized construction of mirror manifolds]. Nucl. Phys. B, 939(1-2):377-391, 1993.

[2] M. Kreuzer and H. Skarke. [On the classification of quasihomogeneous functions]. Commun. Math. Phys., 150(1):137–147, 1992.

Exploring Multi-Colored Peg Solitaire

[1]Melissa Wong, [1]Sonali Vyas, and [2]Alexxis De Lamere
Advisors: [1]Roberto Soto, Ph.D., and [2]Tara Davis, Ph.D.
[1]Department of Mathematics, California State University, Fullerton
[2]Department of Mathematics, Hawaii Pacific University

ABSTRACT

Peg solitaire is a classical one-person game that has been played in various countries on different types of boards. Previous studies have focused on the solvability of the game on these traditional boards; our work focuses on analyzing generalizations of this game and redefining the criteria that determines whether a board is solvable or freely solvable. Specifically, we study the solvability of games with multi-colored pegs on different families of graphs. In conclusion, we arrive at results that differ from previous work on the subject.

Longevity and Immune Defense, Evolutionary Trade-Offs in *Drosophila melanogaster*

Nayelie Y. Benitez Santos and Han Yin

Advisors: [1]Parvin Shahrestani, Ph.D., and [2]Sam Behseta, Ph.D.

[1]Department of Biological Science, California State University, Fullerton

[2]Department of Mathematics, California State University, Fullerton

ABSTRACT

Evolution by natural selection can be constrained by trade-offs between phenotypes. *Drosophila melanogaster*, or the common fruit fly were used in this study to determine the trade-off relationship between immune defense and longevity. Fruit flies' short generation time, ability to produce high numbers of progeny, relatively easy maintenance, and fully sequenced genome, containing many genes homologous to human diseases, make them an ideal species for the investigation of evolutionary trade-offs. We showed that fruit flies experimentally selected for increased fungal resistance for 19 generations had shorter life spans than their controls. After undergoing relaxed selection for nearly 75 generations, the formerly-selected populations maintained their resistance to fungal infection, but no longer paid at a cost to longevity compared to the controls. The disappearance of the trade-off between immune defense and longevity suggests other factors impact immune defense and organismal lifespan. Further studies should be performed to understand this trade-off phenomenon, which could be innovative in the fields of immunology and evolution.

Keywords: *Drosophila melanogaster*, evolutionary trade-offs

INTRODUCTION

Unlike in D. melanogaster, humans have both an innate and adaptive immunity (Croze et al. 2016). The first and only line of defense for D. melanogaster is the innate or nonspecific immune defense in which the host receptors will trigger many defense mechanisms in the presence of any foreign body. Due to similarities in insect and mammalian innate immunity, Drosophila melanogaster has been used as a model to study physiological responses and its relation to immune defense in vertebrates (Buchon et al. 2014). Due to the many similarities in endocrinology that control development, metabolism and immunity, as well as other organ systems to vertebrates, there is relevance in using D. melanogaster for insight into human biology. There are also experimental benefits to using D. melanogaster as a model organism, such as its powerful genetics, low cost, and conserved disease pathways (Pandey & Nichols 2011). Such similarities allow for the use of the fly to make advances in therapies for chronic diseases such as cardiovascular disease and cancer.

D. melanogaster's similarities with other insects also make it a good model system in which to test biocontrol efforts for the possibility of controlling mosquitoes that carry diseases such as malaria and dengue. In this study, Beauveria bassiana was used, an entomopathogenic fungus (Ragavendran et al. 2017). This fungus has biochemicals that can act as a mosquito controlling agent, which can be innovative for reducing the spread of disease, as it is more selective than chemical control.

There is evidence from past research suggesting that immune defense is involved in trade-offs with other life-history traits (Short et al. 2012). The main purpose of this paper is to further look into the relationship between longevity and immune defense in the model organism, D. melanogaster. A study showed that in insects, the nutrients available are a factor that

contribute to the direction of the trade-off in immunity and life span (Ye et al. 2009). In this research, the trade-off of longevity and immune defense was tested through experimental evolution. Improved immune defense evolved at a cost to uninfected longevity. Due to the allocation of resources required for immune defense, it was expected that life-span would decrease. However, once the fruit flies were kept under relaxed selection for 75 generations, it resulted that there was resistance against the fungal infection, but not at the cost of longevity. Such results, support the idea that the immune system is impacted by many factors, and an increased fungal infection resistance can be, but does not have to be, at the expense of life-span.

METHODS AND MATERIALS

Fruit fly population

To allow for there to be genetic diversity amongst the population of fruit flies in this experiment, fruit flies from five geographically different places were outbred. These fruit flies were collected from New York, Zimbabwe, Netherlands, Beijing, and Tasmania, and then the round-robin crossing technique was used on the isofemale lines, such that females of one line were mated with males of another line until all fly lines were mated with each other. In the outbreeding process, the generation cycles were kept at 14-days. To mimic outdoor conditions in the lab, the fruit flies were kept in a 12:12 light/dark incubator. The 80,000 fruit flies that were outbred were separated into eight groups, for a total of 10,000 fruit flies in each group. The treatment condition had a total of 40,000 fruit flies (four groups of 10,000) selected for increased resistance to fungal infection (S). As for the other 40,000 fruit flies, they were kept as controls (C), in four groups of 10,000.

Fungal resistance selection

For each of the four selected (S) populations, in every generation, 10,000 flies were sprayed with 7.5 ml of 0.34 g fungal pathogen B. bassiana suspended in 25 ml 0.03% Silwet using a spray tower. Those flies were separated into groups of 2,000, into five different cages. The flies that were sprayed were kept at 100% humidity, in the dark for the fungus to be able to germinate for 24 hours, which would allow for the cuticles of the flies to be penetrated, therefore infecting the flies. Every day, the flies that died were removed from the cages and fresh food was provided for those that survived. The 20% of the surviving flies that remained, contributed to the next generation by producing offspring. The process was continued for nineteen generations. The control (C) groups had a similar handling process, except there was no fungal infection exposure, and 20% of the flies were randomly selected for them to contribute their offspring to the next generation. To ensure that the following rounds of selection had 10,000 flies, the offspring of each population were put into 100 vials, in which each vial contained 60-80 eggs.

Relaxed Selection

After nineteen generations of fungal resistance selection, the fruit flies were kept under conditions of relaxed selection. For the next seventy-five generations, the flies were not infected with a fungal infection, and were just monitored for survival.

RESULTS AND DATA ANALYSIS

The Kaplan-Meier estimate of the survival function and tests of significance using Cox Proportional Hazard model were performed using the packages Survival and Survminer in software R.

First, we investigated whether or not the replicates of each population contribute significantly to the difference on survival probabilities of *C* and *S* population. Figure 1 plots the curves along with 95% confidence intervals across 4 replicates for two populations. We can see that there is a significant overlap among the confidence intervals of the survival curves of different replicates for each of the *C* and *S* populations. Also, as the pairwise log-rank test (Appendix Table 1) supports the significance of all the confidence interval overlaps,

we decide to treat 4 replicates similar to each other and conclude that the population replicate does not affect the hazard ratio significantly. Similarly, covariate sex doesn't either.

To quantify the effect of other covariates on survival probabilities, we build a Cox proportional hazard model with covariates treatment and population for each of the populations *C* and *S.* However, the proportional hazard assumption is not satisfied: the hazard ratios of treatment and population are changing over time for generation 19, so is the hazard ratio of treatment for generation 94. Thus, we convert them into time-dependent covariates for further modeling. To determine a time-splitting point, we stratify survival curves by treatment and population for generation 19 and by treatment for generation 94 as plotted in Figure 2. Based on the plot and the test of proportional hazard assumption (Appendix Table 2.1 and 2.2), two time-splitting points are chosen at Day 26 and Day 36 for generation 19, and one for generation 94 at Day 36.

With time-dependent covariates included and interaction effects considered, a Cox proportional hazard models is proposed for each of generations 19 and 94 based on likelihood ratio test and the interpretability of a model. At generation 19, the hazard rate h(x,t) of sample *x* at time t is estimated to be,

$$h(x,t)=h_0(t)\ exp\ exp\ [\ I_{inf}(x)\ (\beta_1 I_1+\beta_2 I_2+\beta_3 I_3)+I_{popS}(x)(\beta_4 I_1+\beta_5 I_2+\beta_6 I_3)+\beta_7 I_{inf}(x)^* I_{popS}(x)\ ],$$

where *t* represents the number of days after infection that can be any integer from 0 to 87, $h(x,t)$ represents the hazard rate of sample *x* from generation 19 at time *t*, $h_0(t)$ represents the hazard rate of baseline (uninfected flies of population *C* at generation 19 at time *t*, I_{inf} is 1 if infected and 0 if uninfected, I_{popS} is 1 if population *S* and 0 if population C, I_1 is 1 if the time of interest is between Day 0 and Day 25 and 0 otherwise, I_2 is 1 if the time of interest is between Day 26 and Day 35 and 0 otherwise, I_3 is 1 if the time of interest is between Day 36 and Day 87 and 0 otherwise, $\beta_1, \ldots, \beta_7$ represent the corresponding coefficients, of which the estimations are in.

Also at generation 94, the hazard rate $h(x,t)$ of sample *x* at time t is estimated to be,

$$h(x,t)=h_0(t)\ exp\ exp\ [I_{inf}(x)\ (\beta_1 I_1+\beta_2 I_2)+\beta_3 I_{popS}(x)\ ],$$

where *t* represents the number of days after infection that can be any integer from 0 to 87, $h(x,t)$ represents the hazard rate of sample *x* from generation 94 at time *t*, $h_0(t)$ represents the hazard rate of baseline (uninfected flies of population *C* at generation 60) at time *t*, I_{inf} is 1 if infected and 0 if uninfected, I_{popS} is 1 if population *S* and 0 if population *C*, I_1 is 1 if the time of interest is between Day 0 and Day 35,and 0 otherwise, I_2 is 1 if the time of interest is between Day 36 and Day 87, and 0 otherwise, $\beta_1, \ldots, \beta_7$ represent the corresponding coefficients, of which the estimations are in.

Reading the output of two models (Appendix Table 3.1 and 3.2), we can quantify the differences at certain day after treatment between populations *S* and *C* (baseline) across different treatment status and generations in Table 1.

As seen in Table 1, The Hazard ratio of population *S* against population *C* is changing over time in generation 19, but it stays constant in generation 94. There are different probabilities for generation 19 amongst different days. At any given time between Day 0 to Day 25, the probability of death for an uninfected fly in population *S* is 237% of the probability of death of an uninfected fly in population *C* with p-value 0.00146. Between Day 26 to Day 35, the probability of death for an uninfected fly in population *S* is 182% of the probability of death of an uninfected fly in population *C* with p-value 0.00373. From Day 36 to Day 87, the probability of death for an uninfected fly in population *S* is 131% of the probability of death of an uninfected fly in population *C* with p-value 0.00691. The probability of death for an infected fly in population *S* from Day 0 to Day 25 is 33% of the probability of death of an infected fly in population *C* with p-value < 0.00001. The probability of death for an infected fly in population *S* from Day 26 to Day 35 is 25% of the probability of death

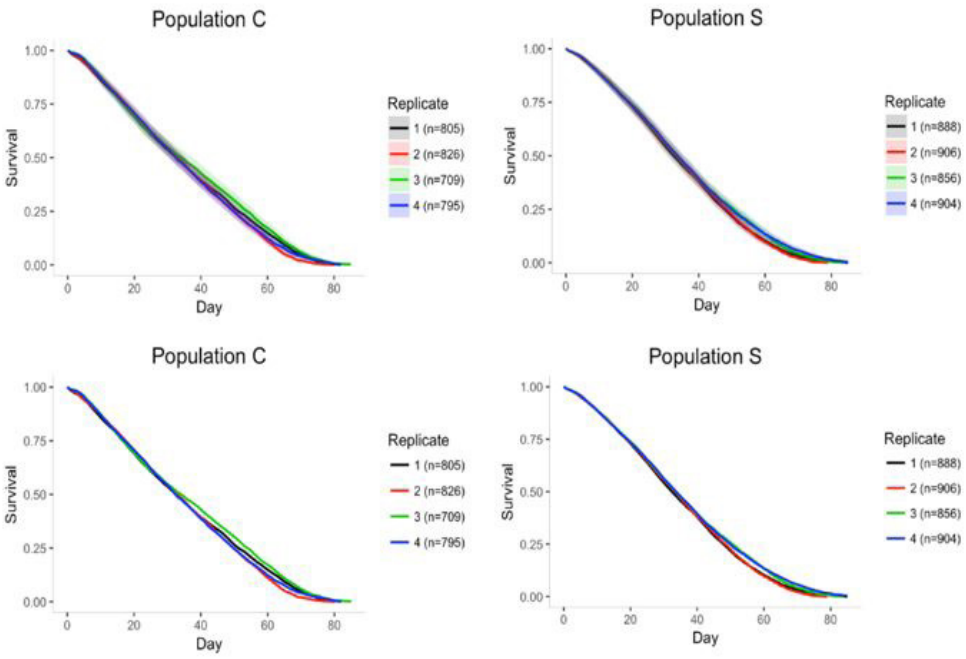


Figure 1. *The Kaplan-Meier estimate of survival functions of population C (left) and population S (right) along with 95% confidence intervals across 4 population replicates.*

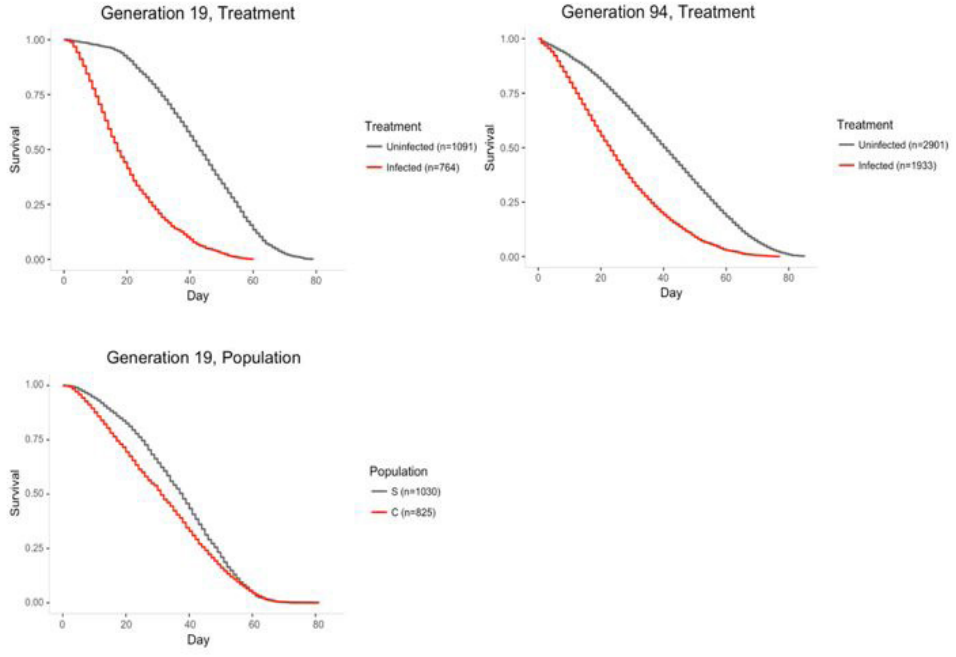


Figure 2. *The Kaplan-Meier estimate of survival functions of generation 19 stratified at treatment (top left) and at population (bottom left), and those of generation 60 stratified at treatment (top right).*

of an infected fly in population *C* with p-value 0.00011. On Day 36 to Day 87, the probability of death for an infected fly in population *S* is 18% of the probability of death of an infected fly in population *C* with p-value 0.00004.

For generation 94, from Day 0 to Day 87, the probability of death for an uninfected fly in population *S* is 92% of the probability of death of an uninfected fly in population *C* with p-value 0.05919. From Day 0 to Day 87 in generation 94, the probability of death for an infected fly in population *S* is 92% of the probability of death of an infected fly in population *C* with p-value 0.05919.

The Cox proportional hazard models also render the predicted survival probabilities of each subgroup at any given time. Figure 3 below visualizes the results along with 95% confidence intervals by bootstrap with 1000 iterations. As the same as the hazard ratios indicate in Table 1, population *S* is more likely to survive compared to population *C* among all infected flies of generation 19, but the relationship is swapped without infection.

In summary, longevity was a trade-off for immunity in the first 19 generations. However, it was found that after 75 of relaxed selection, there was a loss of the trade-off found in previous research. The hazard rate (HR) was evaluated at different points in time for the uninfected and infected populations at generation 19 and 60. For the uninfected flies at days 0-35, the HR is 2.37 (p=0.00146). There is a decrease in the HR on days 36-87, at 1.31(p=0.00691). A decrease in the HR occurs for the infected flies, with a starting HR at 0.33 (p=0.0001), dropping to a 0.18 (0.0004). However, the HR is the same from day 0-35, as it is at day 36-87 for generation 94 at 0.92(p=0.05919). A constant hazard ratio implicates no trade-off, as there is no major die-offs.

Table 1. *The hazard ratios of population S against C (baseline) across different sub-groups by treatment status and the number of days after treatment of generation 19 (top) and of generation 60 (bottom). The p-value of a hypothesis test is reported with the null hypothesis H_o:hazard ratio=1, which means there is no difference between populations S and C in terms of hazard.*

	Generation 19		
	Day 0 – 25	**Day 26 – 35**	**Day 36 – 87**
Uninfected (Baseline C)	2.37 p-value = 0.00146	1.82 p-value = 0.00373	1.31 p-value = 0.00691
Infected (Baseline C)	0.33 p-value < 0.00001	0.25 p-value = 0.00011	0.18 p-value = 0.00004

	Generation 94	
	Day 0 – 35	**Day 36 – 87**
Uninfected (Baseline C)	0.92 p-value = 0.05919	0.92 p-value = 0.05919
Infected (Baseline C)	0.92 p-value = 0.05919	0.92 p-value = 0.05919

DISCUSSION

After 75 generations of relaxed selection, the phenomenon of a trade-off between immune defense and longevity disappeared. This is contrary to results when selecting for improved immune defense, when immune defense came at a cost to longevity resulting in a shorter life span. The results to this research arose a question, in regard to what the cause is of the disappearance of the trade-off. It is possible that a drastic increase in immune defense has a fitness trade-off with longevity, but mild increases in immune defense can be maintained without fitness costs.

Determining the reason(s) of the disappearance in the trade-off could be helpful in determining the impact that the immune system has on lifespan. Continuing such research could be influential in finding medical advances that would influence how the immune system operates, such as therapies that would expand lifespan for humans.

Although sex was not a major focus of this research it could be a variable to look into. Some studies have shown that female fruit flies are more susceptible to infection than male fruit flies, which may be because females allocate resources to produce offspring, which

could be detrimental to their immunity with the possibility that also leads to a shortened lifespan. Other research suggests that mating leads to reduced resistance due to receiving male seminal fluid proteins, Sex Peptide (Schwenke, & Lazzaro, 2017). The juvenile hormone in females after reproduction is an immunosuppressant activated by the Sex Peptide. If such differences can be found to be significant, there could be changes to the way our health system approaches such hardships that reproduction and mating have on the women, possibility creating specialized treatment and medication.

When determining what factors affect longevity, nutrition should be considered. It has been shown that diet, or more specifically the nutrients consumed, influences the lifespan of a fruit fly (Tatar, 2011). The allocations of those nutrients, such as carbon, nitrogen, and amino acids, is also important to consider when determining whether somatic maintenance or reproduction is more important. Finding out the cause of the disappearance of the trade-off between immune defense and longevity, and determining the variables that affect this trade-off would be the next step to this research. Knowing the cause could help in making biological advances, such as reducing disease and better understanding the immune system.

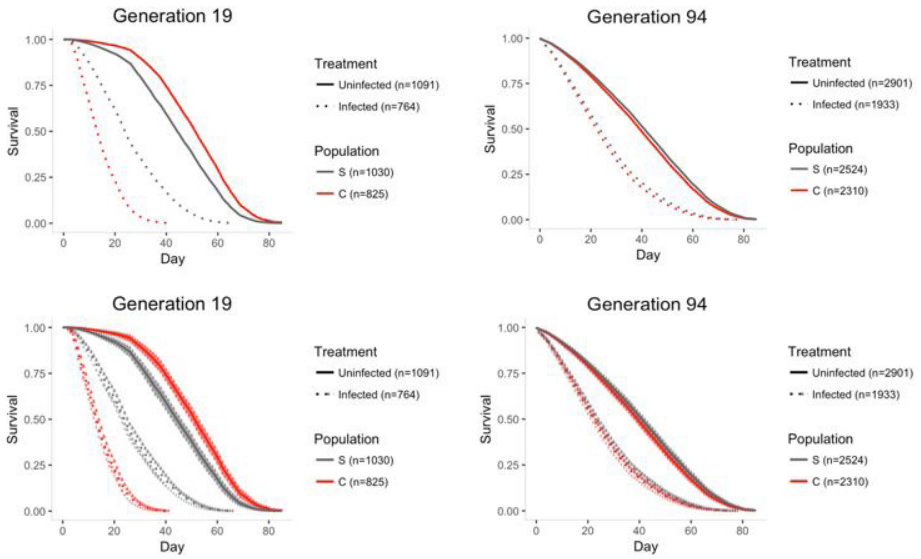


Figure 3. *The predicted survival probabilities by Cox proportional hazard models of different subgroups by populations and treatment status in generation 19 (left) and generation 60 (right). 95% confidence intervals are added (bottom) by bootstrap with 1000 iterations.*

ACKNOWLEDGEMENT

This work was supported by a Big Data Discovery & Diversity – Research Education Advancement Partnership to CSUF from the National Institute of Health [R25MD010397]. We thank students in the Shahrestani Lab who collected the raw data.

REFERENCES

Buchon, N., Silverman, N., & Cherry, S. (2014). Immunity in Drosophila melanogaster -- from microbial recognition to whole-organism physiology. *Nature Reviews. Immunology, 14*(12), 796-810.

Croze, Živković, Stephan, & Hutter. (2016). Balancing selection on immunity genes: Review of the current literature and new analysis in *Drosophila melanogaster*. *Zoology, 119*(4), 322-329.

Garschall, K., & Flatt, T. (2018). The interplay between immunity and aging in *Drosophila. F1000Research, 7*, 160.

Schwenke, & Lazzaro. (2017). Juvenile Hormone Suppresses Resistance to Infection in Mated Female Drosophila melanogaster. *Current Biology,27*(4), 596-601.

Short, Wolfner, & Lazzaro. (2012). Female Drosophila melanogaster suffer reduced defense against infection due to seminal fluid components. *Journal of Insect Physiology, 58*(9), 1192-1201.

Tatar M. (2011). The plate half-full: Status of research on the mechanisms of dietary restriction in *Drosophila melanogaster. Experimental gerontology, 46*(5), 363-368.

Ragavendran, C., Dubey, N., & Natarajan, D. (2017). *Beauveria bassiana* (Clavicipitaceae): a potent fungal agent for controlling mosquito vectors of *Anopheles stephensi, Culex quinquefasciatus* and *Aedes aegypti* (Diptera: Culicidae). *RSC Advances, 7*(7), 3838-3851.

Ye, Y., Chenoweth, S., & Mcgraw, E. (2009). Effective but Costly, Evolved Mechanisms of Defense against a Virulent Opportunistic Pathogen in *Drosophila melanogaster*. *Plos Pathogens, 5*(4), E1000385.

APPENDIX

Table 1. *Pairwise Log-rank Test*

Population Replicates

	1	2	3
2	0.16839	-	-
3	0.16591	0.00024	-
4	0.47547	0.01601	0.47547

P value adjustment method: holm

Table 2. *Proportional Hazard Assumption Test*

Generation 19

	rho	chisq	p
fungal:gt1	-0.01978	0.7198	0.396
fungal:gt2	0.00291	0.0161	0.899
fungal:gt3	0.00964	0.1774	0.674
gt1:popS	-0.00466	0.0414	0.839
gt2:popS	-0.00455	0.0387	0.844
gt3:popS	-0.01960	0.7071	0.400
fungal:popS	-0.00196	0.0073	0.932
GLOBAL	NA	2.9346	0.891

Table 2.2 *Proportional Hazard Assumption Test*

Generation 60

	rho	chisq	p
popS	0.00778	0.293	0.588
fungal:gt1	-0.01487	1.063	0.303
fungal:gt2	-0.01207	0.693	0.405
GLOBAL	NA	1.963	0.580

Table 3.1 *Proportional Hazard Assumption Test*

Generation 19

	coef	exp(coef)	se(coef)	z	Pr(>|z|)	
fungal:gt1	3.75125	42.57433	0.16173	23.195	< 2e-16	***
fungal:gt2	2.87284	17.68718	0.18111	15.862	< 2e-16	***
fungal:gt3	2.94937	19.09396	0.20451	14.422	< 2e-16	***
gt1:popS	0.86201	2.36792	0.17032	5.061	4.17e-07	***
gt2:popS	0.59886	1.82004	0.14043	4.264	2.00e-05	***
gt3:popS	0.26650	1.30539	0.07053	3.778	0.000158	***
fungal:popS	-1.98124	0.13790	0.17509	-11.316	< 2e-16	***

Table 3.2 *Cox Proportional Hazard Model Summary*

Generation 60

	coef	exp(coef)	se(coef)	z	Pr(>|z|)	
popS	-0.08567	0.91790	0.02887	-2.967	0.003	**
fungal:gt1	0.92446	2.52050	0.03880	23.824	<2e-16	***
fungal:gt2	0.60494	1.83115	0.05300	11.413	<2e-16	***

Gone With The Flu

Kristina Rosete, Taylor Salgado, Sonali Vyas
Advisor: Roberto Soto, Ph.D.
Department of Mathematics, California State University, Fullerton

ABSTRACT

Multiple strains of the influenza virus affect people every year, and every year a handful of people get infected with at least one of the many strains. A significant issue in the medical field is how to decrease the population of infected individuals. We created a mathematical model that could determine the rate at which vaccines for the virus would have to be distributed to minimize the quantity of people infected. We accounted for the rate of contracting the influenza virus and the recovery speed. We also considered the possibility of an individual getting infected with more than one strain and restricted the number of vaccines available to make our model more applicable to a real-life situation. This model could help those in the healthcare industry create a more efficient process for vaccinating people and decreasing the number of those infected.

Economical Generating Sets of the Monoid of Partial Order-preserving Injections

Saul Lopez
Advisor: Scott Annin, Ph.D.
Department of Mathematics, California State University, Fullerton

ABSTRACT

Monoids arise in such fields as computer science, physics, and numerous branches of mathematics including abstract algebra, cryptography and operator theory. In this research project we seek to determine economical generating sets for the monoid of partial order-preserving injections of an n-element set, $POI(n)$. A generating set for a monoid is a collection of elements S such that every element of the monoid can be expressed as a product of elements from S. Generating sets are of fundamental importance across math and science, and mathematicians have great interest in studying generating sets of a variety of algebraic structures. By an economical generating set, we refer to a generating set containing as few elements as possible.

Using the algebra software package Groups, Algorithms, and Programming, GAP, we examine the key features of various generating sets of $POI(n)$, including their sizes, the domains and ranges of the partial injections arising in them, and more. We confirm that every generating set of $POI(n)$ contains at least *n* elements, and we describe all economical generating sets (which contain exactly *n* elements). Also, we provide a formula for the number of economical generating sets for $POI(n)$.

INTRODUCTION

Definition 1. A **monoid** is a non-empty set *M* together with a binary operation $*$ that satisfies the following axioms:

1) Closure: If $a, b \in M$, then $a * b \in M$.

2) Associativity: $a * (b * c) = (a * b) * c,$ for all $a, b,$ and $c \in M$.

3) Existence of an Identity: An element $e \in M$ is called the **identity element** if $a * e = a = e * a$ for every $a \in M$.

Example 1. The set of positive real numbers under multiplication, $(\mathbb{R}^+, \times)$, is a classic example of a monoid, with identity element 1.

Note $(\mathbb{R}^+, +)$ is not a monoid since 0 does not belong to $\mathbb{R}^+$.

Example 2. However, $([0, \infty), +)$ is a monoid with identity element 0.

Note that the monoid in Example 2 forms a **submonoid** of $(\mathbb{R}, +)$, since it is a subset of $\mathbb{R}$ that forms a monoid.

Definition 2. A **semigroup** is a non-empty set *S* that satisfies the closure and associativity axioms in Definition 1. Thus, every monoid is a semigroup.

Any semigroup can be made into a monoid by adjoining an identity element. In Example 2, for instance, starting with the semigroup $(\mathbb{R}^+, +)$, we form the monoid $([0, \infty), +) = (\mathbb{R}^+ \cup \{0\}, +)$.

Example 3. Let X be a non-empty set. An expression of the form $x_1x_2x_3 \cdots x_n$ (with $x_i \in X$) is called a **word** on the alphabet X. The set of all words on X is the **free semigroup** on X, denoted X^+. We can form a monoid by adjoining an empty word, say Λ, to X^+. This monoid, denoted X^*, is called the **free monoid** on X.

The operation in X^+ and X^* is concatenation. If we let X be the English alphabet of lowercase letters, then some elements of X^+ are

$$road * runner = roadrunner,$$
$$sledge * hammer = sledgehammer,$$
$$aababba * babbba = aababbababbba.$$

Note that we often omit the operation $*$ from the notation when the operation is known. Thus, for instance, we may write $a * b$ simply as ab.

Example 4. Another example of a monoid is the **bicyclic monoid**, denoted

$$< a, b : \ ab = \Lambda > .$$

The equation $ab = \Lambda$ is called a relation. Starting with the free monoid X^* on $X = \{a, b\}$ we impose this relation to obtain the bicyclic monoid. Using the relation to simplify words, it is easy to see that every element of the bicyclic monoid can be expressed in the form $b^i a^j$ for some $i, j \geq 0$.

Example 5. Let X and Y be any two non-empty sets and define a binary operation on $S = X \times Y$ as follows: $(x_1, y_1) * (x_2, y_2) = (x_1, y_2)$, where $x_1, x_2 \in X$ and $y_1, y_2 \in Y$. Then S forms a semigroup that is commonly known as a rectangular band.

Note that none of the semigroups presented in Examples 3-5 are commutative.

Definition 3. An **idempotent** in a semigroup S is an element $s \in S$ such that $s^2 = s$. A semigroup in which every element is an idempotent is called a **band**.

Note that the semigroup presented in Example 5 is indeed a band.

Example 6. The **symmetric inverse monoid**, ***SIM(S)***, consists of all partial injections from a set *S* to itself. The operation on $SIM(S)$ is given by composition of mappings, where we will adopt the convention that mappings are composed from left to right.

The monoid $SIM(S)$ is of great relevance to this paper. Therefore, let us take some time to elaborate on its structure. Consider the particular case where $S = \{1, 2, 3, \ldots, n\}$, denoted as $SIM(n)$. The identity element of *SIM(n)* is the element that fixes each element of the set $\{1, 2, 3, \ldots, n\}$. The elements of *SIM(n)* can be represented by a classic two-line notation. For example,

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 4 & 6 & - & 5 \end{pmatrix} \in SIM(6). \tag{1}$$

In this element, we observe that 1 maps to 2, 2 maps to 1, 3 maps to 4, 4 maps to 6, 5 does not get mapped, and 6 maps to 5.

Definition 4. The **domain** of an element $\tau \in SIM(S)$ is

$$Dom(\tau) = \{i \in S : \tau(i) \text{ is defined}\}.$$

Definition 5. The **range** of an element $\tau \in SIM(S)$ is

$$Rng(\tau) = \{\tau(i) : i \in S\}.$$

In (1), we can see that $Dom(\tau) = \{1, 2, 3, 4, 6\}$ and $Rng(\tau) = \{1, 2, 4, 5, 6\}$.

Definition 6. The **rank** of an element $\tau \in SIM(S)$ is

$$rank(\tau) = |Dom(\tau)| = |Rng(\tau)|.$$

In (1), $rank(\tau) = 5$. Also, note that the set of all elements in $SIM(n)$ of rank n forms the well-known group called the symmetric group, denoted S_n. Consult any introductory text on modern algebra for more information([6], [8]). To illustrate how composition of partial permutations works, consider this example:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 3 & 4 & - & 5 & 6 & - \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 2 & 1 & 4 & 6 & - & 5 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 4 & 6 & - & - & 5 & - \end{pmatrix}.$$

Of course, the rank of the composition of two partial mappings cannot exceed the rank of either element comprising the product. Here, for example, we see that the elements on the left side have rank 4 and 5, respectively, while the resulting product on the right side only has rank 3.

There are some fundamental equivalence relations which help us understand the structure of semigroups and monoids. These are known as the **Green's relations** [5]. To discuss these, we need some notation: given a monoid M and $t \in M$, denote $Mt = \{mt : m \in M\}$ and $tM = \{tm : m \in M\}$. Now, let $\alpha, \beta \in M$. Then

1. α and β are $\mathcal{L}$**-related**, denoted $\alpha \ \mathcal{L} \ \beta$, if and only if there exist $x, y \in M$ such that $x\alpha = \beta,\ y\beta = \alpha$. That is, $\beta \in M\alpha$ and $\alpha \in M\beta$.
2. α and β are $\mathcal{R}$**-related**, denoted $\alpha \ \mathcal{R} \ \beta$, if and only if there exist $u, v \in M$ such that $\alpha u = \beta,\ \beta v = \alpha$. That is, $\beta \in \alpha M$ and $\alpha \in \beta M$.
3. The $\mathcal{D}$**-relation**, $\mathcal{D} = \mathcal{L} \circ \mathcal{R} = \mathcal{R} \circ \mathcal{L}$, is the smallest equivalence relation containing both $\mathcal{L}$ and $\mathcal{R}$.
4. The $\mathcal{H}$**-relation** is the intersection of $\mathcal{L}$ and $\mathcal{R}$. That is, α and β are $\mathcal{H}$-related, denoted $\alpha \ \mathcal{H} \ \beta$, if and only if $\alpha \ \mathcal{L} \ \beta$ and $\alpha \ \mathcal{R} \ \beta$.

The Green's relations are equivalence relations, and hence form equivalence classes, which we will refer to as $\mathcal{L}$-classes, $\mathcal{R}$-classes, $\mathcal{D}$-classes, and $\mathcal{H}$-classes.

Example 7. Arranging the elements of the bicyclic monoid in an array, we can compute its Green's relations.

$$\begin{array}{cccccc} 1 & a & a^2 & a^3 & a^4 & \dots \\ b & ba & ba^2 & ba^3 & ba^4 & \dots \\ b^2 & b^2a & b^2a^2 & b^2a^3 & b^2a^4 & \dots \\ b^3 & b^3a & b^3a^2 & b^3a^3 & b^3a^4 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{array}$$

Two elements in the bicyclic monoid are $\mathcal{R}$-related if they satisfy the following condition:

$$b^ia^j \ \mathcal{R} \ b^ka^\ell \text{ if and only if } i = k.$$

In terms of the array above, two elements are $\mathcal{R}$-related if and only if they are in the same row. For example, $ba^2 \ \mathcal{R} \ ba^4$, since $ba^2 = ba^4(b^2)$ and $ba^4 = ba^2(a^2)$. Similarly, two elements are $\mathcal{L}$-related if they satisfy

$$b^ia^j \ \mathcal{L} \ b^ka^\ell \text{ if and only if } j = \ell.$$

The elements in the same column of the array above are $\mathcal{L}$-related. For example, $ba^3 \ \mathcal{L}\ b^3a^3$, since $ba^3 = (a^2)b^3a^3$ and $b^3a^3 = (b^2)ba^3$. Next, observe that for all non-negative integers $i, j, k,$ and ℓ,

$$b^i a^j \ \mathcal{D}\ b^k a^\ell$$

since

$$b^i a^j \ \mathcal{R}\ b^i a^\ell \ \mathcal{L}\ b^k a^\ell.$$

Thus, since any two elements of the bicyclic monoid are $\mathcal{D}$-related, the bicyclic monoid has only one $\mathcal{D}$-class. Lastly, it is easy to see that an element is $\mathcal{H}$-related only to itself.

We are now ready to begin studying the particular monoid of interest to us in our research, and in the next section, we will determine its Green's relations.

PARTIAL ORDER-PRESERVING INJECTIONS

The monoid of interest to our research is the monoid of **partial order-preserving injections**, denoted by ***POI(n)***, which is a submonoid of *SIM(n)*. An element σ of $SIM(n)$ is an element of $POI(n)$ if whenever $i < j$ in $\{1, 2, 3, \ldots, n\}$, then $\sigma(i) < \sigma(j)$. For example,

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & - & 5 & 6 & - \end{pmatrix} \in POI(6),$$

since the values on the second row appear in increasing order. However, the element τ in (1) is not an element of $POI(6)$.

Definition 7. A **partial identity** in $POI(n)$ is an element σ such that for each $i \in \{1, 2, 3, \ldots, n\}$, either $\sigma(i) = i$ or $i \notin Dom(\sigma)$.

As we will see in our main results, partial identities play an important role in describing economical generating sets of $POI(n)$. An example of a partial identity is the element

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & - & 3 & - & 5 & - \end{pmatrix} \in POI(6),$$

whose rank is 3. Throughout this paper, we denote the subset of $POI(n)$ of elements of rank k as $POI_k(n)$. For example, $rank(\tau) = 3,$ so $\tau \in POI_3(6)$. Note that $POI(n) = \bigcup_{k=0}^{n} POI_k(n)$, and from [1],

$$|POI(n)| = \sum_{k=0}^{n} \binom{n}{k}^2.$$

The k^{th} term of this summation counts the number of elements in $POI_k(n)$. For example,

$$|POI(4)| = \sum_{k=0}^{4} \binom{4}{k}^2 = \binom{4}{0}^2 + \binom{4}{1}^2 + \binom{4}{3}^2 + \binom{4}{4}^2 = 70.$$

We now look at the Green's relations for the monoid $POI(n)$. Let $\alpha, \beta \in POI(n)$. A routine analysis shows the following:

- $\alpha \; \mathcal{R} \; \beta$ if and only if $Dom(\alpha) = Dom(\beta)$.
- $\alpha \; \mathcal{L} \; \beta$ if and only if $Rng(\alpha) = Rng(\beta)$.
- $\alpha \; \mathcal{D} \; \beta$ if and only if $rank(\alpha) = rank(\beta)$.
- $\alpha \; \mathcal{H} \; \beta$ if and only if $\alpha = \beta$.

We give some examples of the $\mathcal{R}$ and $\mathcal{L}$ relations in $POI(n)$.

Example 8. Let

$$\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 4 & 5 & 6 & 7 & - & - \end{pmatrix} \quad \text{and} \quad \beta = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 3 & 4 & 5 & 6 & - & - \end{pmatrix}$$

in $POI(7)$. Since

$$\alpha \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ - & 2 & - & 3 & 4 & 5 & 6 \end{pmatrix} = \beta$$

and

$$\beta \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ - & 2 & 4 & 5 & 6 & 7 & - \end{pmatrix} = \alpha,$$

we see that $\alpha \; \mathcal{R} \; \beta$.

Example 9. Let

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 4 & 5 & 6 & - & - & 7 & 8 \end{pmatrix} \quad \text{and} \quad \gamma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & - & 4 & 5 & 6 & 7 & 8 & - \end{pmatrix}$$

in $POI(8)$. We have

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & - & 2 & 3 & 4 & 7 & 8 & - \end{pmatrix} \sigma = \gamma$$

and

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 3 & 4 & 5 & - & - & 6 & 7 \end{pmatrix} \gamma = \sigma,$$

so we conclude that $\sigma \; \mathcal{L} \; \gamma$.

Remark 1. Let $\alpha, \beta \in POI_{n-1}(n)$. Then $\alpha\beta \in POI_{n-1}(n)$ if and only if $Rng(\alpha) = Dom(\beta)$. In this case, $\alpha \; \mathcal{R} \; \alpha\beta$ and $\beta \; \mathcal{L} \; \alpha\beta$.

We will now turn our attention to the main problem in this paper, the identification of economical generating sets for $POI(n)$.

Definition 8. A **generating set** for an algebraic structure Z (such as a group, semigroup, monoid, vector space, and so on) is a collection of elements in Z such that each element in Z can be expressed by applying a finite sequence of algebraic operations to those elements.

In linear algebra, a generating set for a vector space is commonly known as a spanning set.

Definition 9. A generating set for Z of smallest possible size is called an **economical generating set**.

In linear algebra, an economical generating set for a vector space V is simply a basis for V. In this case, the size of this economical generating set for V is well-known as the dimension of the vector space, $dim[V]$. A familiar example is the basis $\{i, j, k\}$ for the three-dimensional vector space $\mathbb{R}^3$. The search for economical generating sets of algebraic structures has led to a steady stream of work. In [2], these sets were classified for $SIM(n)$, while in [3], they were studied in the symmetric and alternating groups.

Returning to our study of $POI(n)$, observe that $POI_n(n)$ consists only of the identity element

$$id = \begin{pmatrix} 1 & 2 & 3 & \cdots & n-1 & n \\ 1 & 2 & 3 & \cdots & n-1 & n \end{pmatrix}.$$

This element is uninteresting with respect to the generating process, so we will focus our attention on generating the set $POI(n) - \{id\}$. For the remainder of this paper, when we speak of generating $POI(n)$, we will mean that we are generating $POI(n) - \{id\}$, which itself is a subsemigroup of $POI(n)$. The key to generating $POI(n) - \{id\}$ is to generate all of the rank $n-1$ elements first. In fact, it is customary to restrict generating sets of $POI(n)$ to consist of elements of rank $n-1$. By multiplying elements of $POI_{n-1}(n)$ together, we can obtain elements of lower rank (see Remark 1). We arrange the elements in $POI_{n-1}(n)$ in an $n \times n$ array as shown in Figure 1 for $POI(4)$. Elements in the same row have the same domain, elements in the same column have the same range, and elements on the main diagonal are partial identities. Let

$$\widehat{S}_i = \{1, 2, \ldots, i-1, i+1, \ldots, n-1, n\}$$

for $1 \leq i \leq n$. For example, for $n = 4$, if

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & - & 4 \end{pmatrix}, \tag{2}$$

then $Dom(\sigma) = \widehat{S}_3$ and $Rng(\sigma) = \widehat{S}_1$.

	Column 1	*Column* 2	*Column* 3	*Column* 4	
Row 1	$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & - \end{pmatrix}$	$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & - \end{pmatrix}$	$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & - \end{pmatrix}$	$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & - \end{pmatrix}$	$\widehat{S}_4$
Row 2	$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & - & 3 \end{pmatrix}$	$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & - & 4 \end{pmatrix}$	$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & - & 4 \end{pmatrix}$	$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & - & 4 \end{pmatrix}$	$\widehat{S}_3$
Row 3	$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & - & 2 & 3 \end{pmatrix}$	$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & - & 2 & 4 \end{pmatrix}$	$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & - & 3 & 4 \end{pmatrix}$	$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & - & 3 & 4 \end{pmatrix}$	$\widehat{S}_2$
Row 4	$\begin{pmatrix} 1 & 2 & 3 & 4 \\ - & 1 & 2 & 3 \end{pmatrix}$	$\begin{pmatrix} 1 & 2 & 3 & 4 \\ - & 1 & 2 & 4 \end{pmatrix}$	$\begin{pmatrix} 1 & 2 & 3 & 4 \\ - & 1 & 3 & 4 \end{pmatrix}$	$\begin{pmatrix} 1 & 2 & 3 & 4 \\ - & 2 & 3 & 4 \end{pmatrix}$	$\widehat{S}_1$
	$\widehat{S}_4$	$\widehat{S}_3$	$\widehat{S}_2$	$\widehat{S}_1$	

Figure 1: 4×4 *array of the elements in* $POI_3(4)$.

We can represent the elements in Figure 1 by their position in the array as we have done in Figure 2. This will let us denote an element σ in $POI_{n-1}(n)$ as a pair of numbers, (a, b), according to the position of σ in row a and column b of the array. The element σ in (2) can now be represented as $\sigma = (2, 4)$. Note that if $Dom(\sigma) = \widehat{S}_i$ and $Rng(\sigma) = \widehat{S_j}$, then we have $\sigma = (n - i + 1, n - j + 1)$.

$$\begin{array}{cccc} (1,1) & (1,2) & (1,3) & (1,4) \\ (2,1) & (2,2) & (2,3) & (2,4) \\ (3,1) & (3,2) & (3,3) & (3,4) \\ (4,1) & (4,2) & (4,3) & (4,4) \end{array}$$

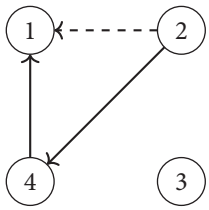
Figure 2: The elements in $POI_3(4)$ represented by their position in the 4×4 array in Figure 1.

Notation: Let X be a subset of $POI_{n-1}(n)$. Let Ω_X denote a directed graph on $\{1, 2, \ldots, n\}$, where a $\frown$ b is a directed edge in Ω_X if and only if $(a, b) \in X$.
If $\sigma = (a, b)$ and $\tau = (c, d)$ are two elements of $POI_{n-1}(n)$, then note that $\sigma\tau = (a, b)(c, d) = (a, d)$ if $b = c$. If $b \neq c$, by Remark 1 the rank of $\sigma\tau$ drops and we do not express $\sigma\tau$ as a pair. Multiplication can now be performed by following directed edges, as shown in our next examples.

Example 10. In $POI(4)$,

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & - & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ - & 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & - & 3 \end{pmatrix}$$

can be expressed as $(2, 4)(4, 1) = (2, 1)$, which is represented graphically as



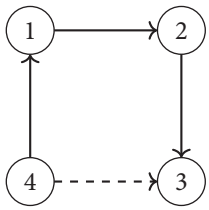
(3)

In (3), the dotted arrow indicates the result of the composition of the two solid arrows.

Example 11. Again in $POI(4)$,

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ - & 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & - \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & - & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ - & 1 & 3 & 4 \end{pmatrix}$$

can be represented as $(4, 1)(1, 2)(2, 3) = (4, 3)$, which is represented graphically as



(4)

In (3) and (4), the solid directed edges form an example of a *walk* [4].

Definition 10. A **walk** is a sequence of m edges of the form

$$x_0 \to x_1 \to x_2 \to \cdots \to x_m,$$

where $x_0, x_1, x_2, \ldots, x_m$ are (not necessarily distinct) vertices, and $x_i \to x_j$ indicates that vertices x_i and x_j are joined by an edge. If $x_0 = x_m$, then the walk is closed.

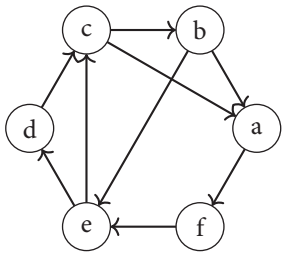


Figure 3: An example of a directed graph.

For example, in Figure 3 we have a walk $c \to b \to e \to d \to c \to a$. Note that if we remove the edge $c \to a$, we have a closed walk: $c \to b \to e \to d \to c$.

Definition 11. If a walk is composed of distinct edges and has distinct vertices, then it is called a **path**. That is, in a path we have $x_i \neq x_j$ for $0 \leq i < j \leq m$, except possibly for $x_0 = x_m$. In the case where $x_0 = x_m$, the path is closed and usually referred to as a **cycle** or *m*-**cycle**.

In Figure 3, we have the path $e \to d \to c \to b$ and the 4-cycle $e \to d \to c \to b \to e$.

Definition 12. A **Hamiltonian path**, also called a Hamilton path, is a path between two vertices of a graph that visits each vertex of the graph exactly once. A closed Hamiltonian path is a **Hamiltonian cycle**.

In Figure 3, we have the Hamiltonian path $b \to e \to d \to c \to a \to f$, and a Hamiltonian cycle $e \to d \to c \to b \to a \to f \to e$.

Shortly, we will see how this basic graph theory can help us describe economical generating sets of *POI(n)*. Our results were first motivated by a program written using the programming language Groups, Algorithms, and Programming (GAP) [7]. The program is explicitly provided for the case $n = 3$ in Figure 4. As we will see in Section 3, any economical generating set of $POI(n)$ must contain exactly n elements. Therefore, the program takes every collection of n elements from the n^2 elements of $POI_{n-1}(n)$ and computes the submonoid of $POI(n)$ that they generate to determine whether or not each such collection is in fact a generating set.

```
gap> basisElements:= function(n)
> local m,i,j,k,y,count;
> count:=0;
> y:=[];
> for i in [2..n^2+1] do
> for j in [i+1..n^2+1] do
> for k in [j+1..n^2+1] do
> count:= count+1;
> y[count]:= Monoid(List(POI(n))[i], List(POI(n))[j], List(POI(n))[k]);
> if y[count]= POI(n) then
> Print("[", List(POI(n))[i], List(POI(n))[j], List(POI(n))[k], "]", "\n");
> fi;
> od;
> od;
> od;
> end;
```

Figure 4: Program code for $POI(3)$, which can easily be modified for $n > 3$. For $n \geq 7$, the number of elements in $POI(n)$ becomes too large for this program to compile.

Using the program, we obtained the following:

Table 1. *Number of economical generating sets of $POI(n)$.*

n	Number of economical generating sets of $POI(n)$
3	2
4	6
5	24
6	120

We may suspect from Table 1 that the number of economical generating sets of $POI(n)$ is $(n-1)!$, and later in this paper, we will prove that this is indeed true.

MAIN THEOREMS

We now turn to the main results in this paper. Our first result provides a necessary condition for a subset $X \subseteq POI_{n-1}(n)$ to be an economical generating set for $POI(n)$.

Theorem 1. Let X be a subset of $POI_{n-1}(n)$ consisting of n elements that generate $POI(n)$. Then X has no partial identities, no $\mathcal{R}$-related elements, and no $\mathcal{L}$-related elements.

Proof. We proceed by way of contradiction. First assume that there are two $\mathcal{R}$-related elements in X. Then for some $i \in \{1, 2, \ldots, n\}$, the subset $\widehat{S}_i$ is the domain for no elements in X. Let $\gamma \in POI_{n-1}(n)$ be any element such that $Dom(\gamma) = \widehat{S}_i$. Write $\gamma = (a, b)$, where we know that $n - i + 1 = a,$ or $i = n - a + 1$. Since X is a generating set, then there exist $x_1, x_2, \ldots, x_j \in X$ such that

$$\gamma = x_1 x_2 \cdots x_j.$$

Writing $x_i = (x_{i1}, x_{i2})$ for each i with $1 \leq i \leq j,$ we have

$$\gamma = (x_{11}, x_{12})(x_{21}, x_{22}) \cdots (x_{j1}, x_{j2}) = (a, b),$$

where $x_{i2} = x_{i+1,1}$ for each i with $1 \leq i < j$. It follows that $a = x_{11}$, and thus, $Dom((x_{11}, x_{12})) = \widehat{S}_i$ which is a contradiction.

Similarly, assume X contains two $\mathcal{L}$-related elements. Thus, there is some subset $\widehat{S}_i$ that no element in X has as its range. Let $\alpha = (a, b) \in POI_{n-1}(n)$, where $Rng(\alpha) = \widehat{S}_i$. Then there exist $x_1, x_2, \ldots, x_j \in X$ such that $\alpha = x_1 x_2 \cdots x_j$. Writing $x_i = (x_{i1}, x_{i2})$ as before, we have

$$\alpha = (x_{11}, x_{12})(x_{21}, x_{22}) \cdots (x_{j1}, x_{j2}) = (a, b),$$

where $x_{i2} = x_{i+1,1}$ for each $1 \leq i < j$. It follows that $b = x_{j2}$, which implies $Rng((x_{j1}, x_{j2})) = \widehat{S}_i$, which is a contradiction.

Now assume X contains a partial identity element, say $\omega = (a, a)$, which fixes the elements of $\widehat{S}_{n-a+1}$. Choose any element $\gamma = (a, b) \in POI_{n-1}(n)$ with $\gamma \neq \omega$. Then, following the notation above, γ can be expressed as

$$\gamma = x_1 x_2 \cdots x_j = (x_{11}, x_{12})(x_{21}, x_{22}) \cdots (x_{j1}, x_{j2}) = (a, b).$$

As above, we get $a = x_{11}$. Thus, since the only element of X of the form (a, y) is the partial identity (a, a) it must be the case that $(x_{11}, x_{12}) = \omega$. Hence,

$$\gamma = (a, a)(x_{21}, x_{22}) \cdots (x_{j1}, x_{j2}).$$

Now we deduce that $x_{21} = a$ and we can repeat the same reasoning as above to conclude that $x_{22} = a$. Therefore,

$$\gamma = (a, a)(a, a) \cdots (x_{j1}, x_{j2}).$$

Proceeding in this way, we eventually have

$$\gamma = (a, a)(a, a) \cdots (a, a) = (a, a) = \omega,$$

a contradiction. □

As a consequence of the proof of Theorem 1, observe that every $\widehat{S}_i$ (for $1 \leq i \leq n$) must be the domain of some generator in a generating set $POI(n)$. This therefore proves the following:

Corollary 1. Every generating set of $POI(n)$ must contain at least n elements.

Theorem 2. Let X be a subset of n elements of $POI_{n-1}(n)$ with no partial identities, no $\mathcal{R}$-related elements, and no $\mathcal{L}$-related elements. Then X is a generating set for $POI(n)$ if and only if the edges of Ω_X form an n-cycle.

Proof. First, assume X is a generating set for $POI(n)$. By way of contradiction, suppose that the edges of Ω_X do not form an n-cycle. This implies we must have at least two smaller cycles, say C_1 and C_2, of lengths n_1 and n_2, respectively. Let us choose an element $(a, b) \in POI_{n-1}(n)$, where a is a vertex of the n_1-cycle C_1 and b is a vertex of the n_2-cycle C_2. Since X is a generating set, we can express (a, b) as a product of generators:

$$(a, b) = (x_{11}, x_{12})(x_{21}, x_{22}) \cdots (x_{j1}, x_{j2}).$$

However, since $x_{11} = a$ is a vertex in C_1, then x_{12} is also a vertex in C_1, which implies $(x_{11}, x_{12}) \in \Omega_X - C_2$. Similarly, since $x_{12} = x_{21}$ is a vertex in C_1 then $x_{22} = x_{31}$ is a vertex in C_1 which implies $(x_{21}, x_{22}) \in \Omega_X - C_2$. Continuing this process we get $(x_{j1}, x_{j2}) \in \Omega_X - C_2$which is a contradiction to the fact that $x_{j2} = b$ is a vertex in C_2. Hence, Ω_X forms an n-cycle.

Now assume that the edges of Ω_X form an n-cycle. To show that X is a generating set, let (a, b) be an arbitrary element in $POI_{n-1}(n)$. Since the edges of Ω_X form an n-cycle, then we have a Hamiltonian cycle of the form

$$x_1 \to x_2 \to \cdots \to x_{i-1} \to a \to x_{i+1} \to \cdots \to x_{j-1} \to b \to x_{j+1} \to \cdots \to x_n \to x_1$$

or

$$x_1 \to x_2 \to \cdots \to x_{i-1} \to b \to x_{i+1} \to \cdots \to x_{j-1} \to a \to x_{j+1} \to \cdots \to x_n \to x_1$$

In the former case, there is a path connecting a and b namely

$$a \to x_{i+1} \to \cdots \to x_{j-1} \to b.$$

(The later case is similar.) Therefore, (a, b) can be generated by X as

$$(a, b) = (x_i, x_{i+1})(x_{i+1}, x_{i+2}) \cdots (x_{j-1}, x_j).$$

Once all the elements of $POI_{n-1}(n) - \{id\}$ are generated, it is trivial to generate the lower rank elements of $POI(n) - \{id\}$.

□

Example 12. The set

$$X = \left\{ \underset{(4,1)}{\begin{pmatrix} 1 & 2 & 3 & 4 \\ - & 1 & 2 & 3 \end{pmatrix}}, \underset{(1,2)}{\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & - \end{pmatrix}}, \underset{(2,3)}{\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & - & 4 \end{pmatrix}}, \underset{(3,4)}{\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & - & 3 & 4 \end{pmatrix}} \right\}$$

forms a generating set for $POI(4)$ since the edges of the graph Ω_X form a 4-cycle. By composing these edges, we can then generate all 16 edges of $POI_3(4)$ to get:

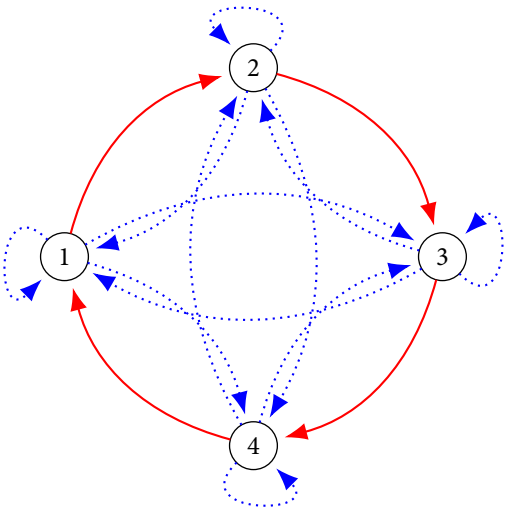


Figure 5: A graphical example of a generating set of $POI(4)$.

Here, the solid directed arrows belong to the generating set, while the dotted directed arrows represent the remaining elements of $POI_3(4)$.

Example 13. The set

$$X = \left\{ \underset{(1,3)}{\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 4 & 5 & - \end{pmatrix}}, \underset{(2,1)}{\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & - & 4 \end{pmatrix}}, \underset{(3,5)}{\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & - & 4 & 5 \end{pmatrix}}, \underset{(4,2)}{\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & - & 2 & 3 & 5 \end{pmatrix}}, \underset{(5,4)}{\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ - & 1 & 3 & 4 & 5 \end{pmatrix}} \right\}$$

forms a generating set for $POI(5)$ since the edges of the graph Ω_X form a 5-cycle. From this 5-cycle, we can generate the remaining 20 edges that correspond to the remaining 20 elements of $POI_4(5)$:

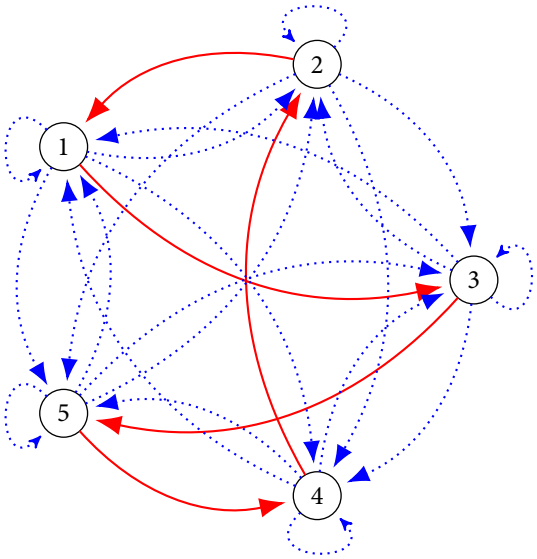


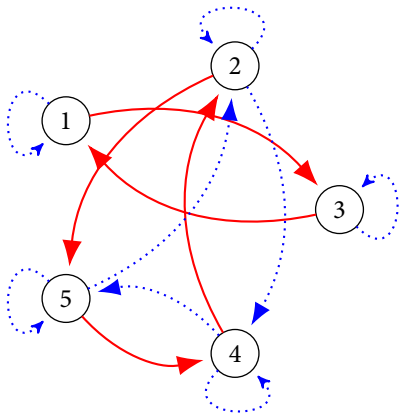
Figure 6: A graphical representation of a generating set of $POI(5)$.

As in Example 12, the solid directed arrows belong to the generating set, while the dotted directed arrows represent the remaining elements of $POI_4(5)$.

Example 14. The set

$$X = \left\{ \underset{(5,1)}{\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & - & 3 & 4 \end{pmatrix}}, \underset{(3,2)}{\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & - & 2 & 3 & 5 \end{pmatrix}}, \underset{(2,3)}{\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 4 & 5 & - \end{pmatrix}}, \underset{(1,4)}{\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ - & 1 & 3 & 4 & 5 \end{pmatrix}}, \underset{(4,5)}{\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & - & 5 \end{pmatrix}} \right\}$$

is not a generating set for $POI(5)$ since Ω_X forms two subcycles of the 5-cycle. By composing the edges in these subcycles we only generate 13 out of the 25 edges of $POI_4(5)$:



.

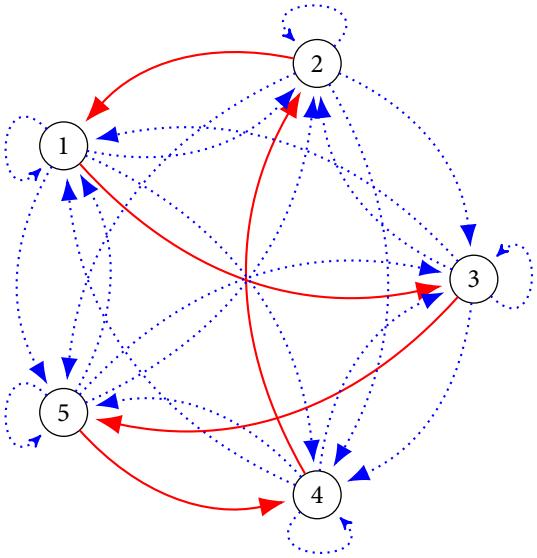


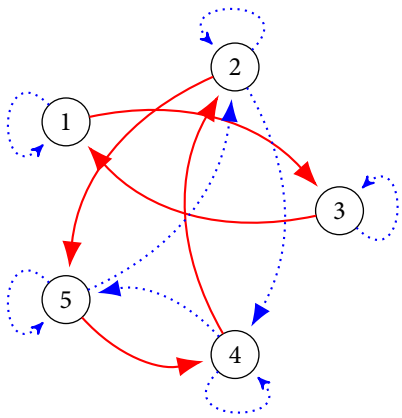
Figure 6: A graphical representation of a generating set of $POI(5)$.

As in Example 12, the solid directed arrows belong to the generating set, while the dotted directed arrows represent the remaining elements of $POI_4(5)$.

Example 14. The set

$$X = \left\{ \underset{(5,1)}{\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & - & 3 & 4 \end{pmatrix}}, \underset{(3,2)}{\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & - & 2 & 3 & 5 \end{pmatrix}}, \underset{(2,3)}{\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 4 & 5 & - \end{pmatrix}}, \underset{(1,4)}{\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ - & 1 & 3 & 4 & 5 \end{pmatrix}}, \underset{(4,5)}{\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & - & 5 \end{pmatrix}} \right\}$$

is not a generating set for $POI(5)$ since Ω_X forms two subcycles of the 5-cycle. By composing the edges in these subcycles we only generate 13 out of the 25 edges of $POI_4(5)$:



.

Theorem 3. *POI(n)* has $(n-1)!$ economical generating sets.

Proof. From Theorem 2, we see that the number of economical generating sets of $POI(n)$ is equal to the number of different n-cycles on a graph with n vertices. It is a well known fact from combinatorics that the number of such n-cycles is

$$\frac{n!}{n} = (n-1)!$$

□

REFERENCES

[1] Alvarez, U. (2015) On kth roots in semigroups of order-preserving partial permutations. CSUF NSM Dimensions, 123-132.

[2] Annin, S. A. (1997) Hierarchy of efficient generators of the symmetric inverse monoid. Semigroup Forum, 55, 327- 355.

[3] Annin, S. A. and Maglione, J (2012) Economical generating sets for the symmetric and alternating groups consisting of cycles of a fixed length. Journal of Algebra and its Applications, Vol. 11, No.6.

[4] Brualdi, R. A., Introductory Combinatorics. 5th ed., Pearson/Prentice Hall, Upper Saddle River, NJ, 2010.

[5] Clifford, A. H. and Preston, G. B., The Algebraic Theory of Semigroups. Vol. 1, American Mathematical Society, Providence, RI, 1961.

[6] Dummit D. S. and Foote R. M., Abstract Algebra 3th ed., John Wiley and Sons Inc, 2004.

[7] GAP 2017, The GAP Group, GAP-Groups, Algorithms, and Programming, Version 4.8.7: 2017. (HTTP://WWW.gap-system.org)

[8] Hungerford, W. T., Abstract Algebra: An Introduction 3th ed., Brooks/Cole, Cengage Learning, 2013.

Student Blending in Math and Physics Integration Problems

Anthony Pina
Advisor: Michael Loverude, Ph.D.
Department of Physics, California State University, Fullerton

ABSTRACT

Introductory calculus-based physics courses present situations in which students are asked to apply tools acquired in math courses to physical phenomenon. For this poster we look specifically at integration and how students use it to solve problems involving variable forces in two different courses: introductory calculus-based mechanics and second-semester calculus. These courses are often taken concurrently and require students to blend information from both disciplines to create a full picture of the solution to physics problems. Data were collected using semi-structured think-aloud interviews including math and physics contexts. Data were analyzed using a cognitive blending lens [Fauconnier & Turner 2002]. While these tasks would seem to prompt blending, students struggled with constructing an effective blend.

The Stochastic Force Spectrum of a Micro-Swimmer

Corbyn Jones and Mauricio Gomez
Advisor: Wylie Ahmed, Ph.D.
Department of Physics, California State University, Fullerton

ABSTRACT

We seek to understand the dynamics of micro-swimmers by quantifying the stochastic forces generated by their motion. We are currently working with Chlamydomonas Reinhardtii—a green algae commonly used to study microscopic locomotion. Our approach is to use optical tweezers and a direct force calibration known as the photon momentum method (PMM) to measure micro-swimmer forces. The power spectral density (PSD) of the force dynamics is analyzed, providing information about the frequency content of the force signals. A simple stochastic model based on the generalized Langevin equation predicts the power spectral density to have a Lorentzian-type curvature. We compare our experimental data to the theoretical model to test if the model can predict our experimentally measured PSD. This approach allows the calculation of thermodynamic quantities such as work, power, efficiency, etc. to describe the microscopic motion. Our analysis seeks to apply concepts from stochastic thermodynamics to understand micro-swimmer dynamics.

Building a Custom Microscope to Study Brownian Motion and Active Matter

Hunter Seyforth
Advisor: Wylie Ahmed, Ph.D.
Department of Physics, California State University, Fullerton

ABSTRACT

Our goal is to build an optical microscope, calibrate it, and make precise measurements of Brownian motion and diffusion using multiple approaches such as mean squared displacement (MSD) analysis and differential dynamic microscopy (DDM). These methods of analysis were applied to quantify the motility of active matter and standardize the process to develop an advanced module for the graduate program. We constructed an optical microscope based on the design by Kemp et al.(arXiv:1606.03052). Then, with a 40x objective, the Brownian motion of 1 micron colloidal particles were studied and used MATLAB to calculate the diffusion coefficient of the particles using two independent approaches. Both single particle tracking and image correlation techniques were implemented to analyze colloidal diffusion. To do this, publicly available matlab codes for particle tracking, MSD analysis, and DDM analysis were applied to calculate the diffusion coefficient. These analysis methods were then utilized to quantify the diffusion of active matter. Specifically, we focused on the diffusion of two different types of active matter: janus particles and swimming microorganisms. This was done to quantify the motility of active matter and compare its diffusion to non-active matter. This project is being developed into an advanced lab module to be an introduction to physics research, fortify concepts from optics and statistical physics, and give students hands-on experience in building optical systems and analyzing noisy data.

Optical Tweezer Measurements in Chlamydomonas

[1]Mauricio Gomez and [1,2]Corbyn Jones
Advisor: [1]Wylie Ahmed, Ph.D.
[1] Department of Physics, California State University, Fullerton
[2] Department of Mechanical Engineering, California State University, Fullerton

ABSTRACT

Optical tweezers have been used to study force fluctuations in microscopic systems and to determine the physical properties of complex materials. We implement two calibration methods, the photon-momentum method and the active-passive method, to measure displacements and forces at the nanometer and piconewton scales. Here, we study the force fluctuations of chlamydomonas microswimmers by applying the photon-momentum method. We calculated the stochastic force spectrum (P_{total}) by estimating the power spectral density of the fluctuating force signal. We apply a theoretical framework to extract the non-equilibrium microswimmer forces from the total force spectrum. We then apply a spectrogram analysis to understand the frequency-time dynamics of the chlamydomonas microswimmer. We hope to extract the force kinetics of the flagella and molecular motors.

Vortex and Tendex Line Simulations of Head on Merging Black Holes

Samuel Rodriguez
Advisor: Geoffrey Lovelace, Ph.D.

ABSTRACT

The Weyl curvature tensor $C_{\alpha\beta\gamma\delta}$ can be split into two parts, the electric and magnetic part. The electric part ε_{jk} describes tidal gravity (how much objects are stretched and squeezed) and the magnetic part B_{jk} describes differential frame dragging (how much objects are twisted) of spacetime. Vortex lines are tangent to the eigenvectors of B_{jk} and tendex lines are tangent to the eigenvectors of ε_{jk}. These lines are analogous to electric and magnetic field lines. We present visualizations of vortex and tendex lines of merging black holes colliding head on. We attempt to make some discussions on how warped spacetimes behave around binary black holes by beginning on the most trivial cases and adding in parameters such as spin and orbital eccentricity. This is the first time we have been able to track vortex and tendex consistently as the simulation progresses.

Authors and Editors

Alejandra Garcia

Alejandra Garcia is a senior Louis Stokes Alliance for Minority Participation (LSAMP) scholar at California State University, Fullerton (CSUF) pursuing a Bachelor of Science Degree in Biology. She works with Dr. Maria Linder purifying and studying copper carrying components of human and pig plasma. Upon graduation, she plans to continue her education as either an M.D., or MD – PhD.

Alexxis De Lamere

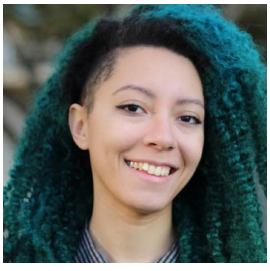
Alexxis De Lamere is a honors undergraduate at Hawaii Pacific University. She double majors in oceanography and applied mathematics,t and plans to be an oceanographic engineer that focuses on sustainability in costal areas around the world. She is currently working on a research project with other undergraduate students from Cal State Fullerton and they plan to publish a paper on their work by the end of Summer 2019.

Andrew Jaramillo

Andrew is a senior Biology major with a concentration in ecological and evolutionary biology. He is a scholar of the Southern California Ecosystems Research Program (SCERP) and is working under Dr. Kristy Forsgren to understand the comparative reproductive anatomy of male surfperches. After his graduation this spring he plans to work in wildlife biology focusing on restoration of our natural wildlife.

Ariana Sanchez

Ariana Sanchez is majoring in Biological Sciences at California State University, Fullerton. She has been a Southern California Ecosystems Research Program (SCERP) scholar since 2018. Ariana is currently working with Dr. Danielle C. Zacherl to quality the health of filter feeding marine invertebrates in an oyster restoration study using condition index and biomass metrics. After graduating in Spring 2019, she plans to gain additional experience through a USGS internship and further her career in ecological research.



Brittany Cook

Brittany Cook is a 2017-2019 Scholar for the Southern California Ecosystems Research Program (SCERP) at California State University, Fullerton. She has a passionate love for plants and algae and has been involved in independent research ever since joining SCERP. She hopes to pursue a Ph.D. in agricultural science after graduating to increase the efficiently, affordability, and sustainability of crop plants for the world's growing population.

Cameron Marzvaan

Cameron Marzvaan is a graduate of the University of California, Santa Barbara (UCSB) with a Bachelor of Science Degree in Biopsychology. Currently, he contributes to research in Dr. Maria Linder's lab using a variety of experimental techniques to purify and study copper carrying proteins in human and swine. As an aspiring physician, he enjoys community service and is applying to medical school this summer.

Christian Concha

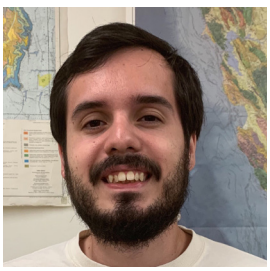
Christian Concha is a student at California State University Fullerton expected to graduate summer 2019. Christian is working with Diane Clemens-Knott using U-Pb Dating to determine the Maximum Depositional Ages and Provenance of the Cow Creek and Fairview metasedimentary Pendants, of the southern Sierra Nevada Mountains. He was recently awarded the Margaret Skillman Woyski Scholarship for his research efforts and seeks to obtain a career in Environmental or Geophysical consulting.

Corbyn Jones

Corbyn Jones is a senior at Cal State Fullerton, majoring in mechanical engineering and physics with an emphasis in business. He plans to graduate in May 2020 and enroll in a graduate program to earn a doctorate in mechanical engineering. He hopes to launch his own company specializing in cutting edge, robotic prostheses for amputees and recently presented his biophysics research at the March 2019 American Physical Society (APS) Conference in Boston.

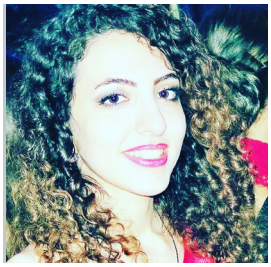
Crystal Paulus

Crystal is a fourth-year undergraduate majoring in chemistry. She has been in Dr. de Lijser's organic chemistry lab since Fall 2017, working on understanding the structure and reactivity of reactive intermediates. After graduating this spring, Crystal will attend Boise State University, pursuing a master's degree in chemistry. Overall her goal is to pursue a career in the cosmetic industry.



Eddie Reyes

Eddie Reyes is an undergraduate geology major. He is completing an undergraduate thesis with Dr. Knott on Discharge and Drought Recovery of Three Springs in Harper Valley, Mojave Desert, California. He is active in the geology club and plans to work in the environmental consulting industry in the fall. He plans to return to graduate school following a few years of work.



Elnaz Bagheri

Elnaz Bagheri is pursuing a Master's degree in Biology at CSUF. Bagheri was a research assistant at Molecular Medicine Center, Rafsanjan Iran for two years before starting her master's degree. She passed 60 hours of courses in Molecular and Genetic engineering at Shahid Beheshti University in Tehran and obtained a Certificate from the Iran Biotechnology Society and Gene Transfer Pioneers Research Group (GTP) in 2015. She presented her poster on March 22,2018 at NSM Symposium at CSUF.

Freddy Nungaray

Freddy is completing his final year at CSUF as an applied mathematics major. Freddy was involved in the Graduate Readiness and Access in Mathematics (GRAM) to enhance his professional growth. Upon completing GRAM, Freddy also conducted research in taxicab geometry over at Seattle University. His interests include research involving geometry, numerical analysis, and mathematical modeling. Freddy intends to pursue a Master's in applied mathematics and aspires to do industrial work.

Gabriel Martinez

Gabriel earned his Associate of Science in Mathematics degree with Honors from Santiago Canyon College, and then transferred to CSUF. In addition to his classes, Gabriel has worked as a supplemental instruction leader for calculus courses, and he has also tutored K-12 students in Common Core mathematics. Gabriel has worked under the mentorship of applied mathematician, Dr. Laura S. Chowdhury, for the Graduate Readiness and Access in Mathematics program (GRAM).

Han Yin

Han Yin is currently a graduate student majoring in Statistics at California State University, Fullerton. Mentored by Dr. Sam Behseta and Dr. Parvin Shahrestani, she applies statistical techniques on the survival data of flies. She will graduate in 2020 and her future goal is to pursue a PhD in Statistics.

Holly Suther

Holly is pursuing a Bachelor of Science degree in Biological Science with a concentration in Marine Biology. She is a Southern California Ecosystems Research Program (SCERP) scholar and works under the guidance of Dr. Kristy Forsgren describing the external morphology of the urogenital papilla on rockfish. Holly is a Supplemental Instruction (SI) Leader for Biology and the Co-President of the Society for Advancement of Chicanos/ Hispanics and Native Americans in Science (SACNAS) organization on campus.

Jackson Flanagan

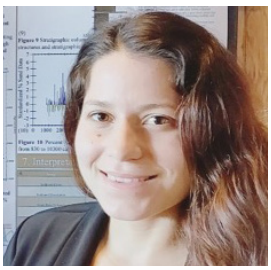
Jackson Flanagan has learned and experienced so much since he transferred into CSUF from Santiago Canyon College as a Geology major two years ago. As an undergraduate researcher at CSUF, he focused on constraining the orientation of the Sierra Nevada Frontal Fault System. Now that he is graduating, he is looking to get into the environmental consulting industry and eventually becoming a Professional Geologist

Jamie Hayward

Jamie Hayward is pursuing a Bachelor of Science in Geology at California State University, Fullerton. She works under the guidance of Dr. Valbone Memeti researching textual and compositional variability of magma bodies from the mid-crust, exhumed at the surface. She is the current social media coordinator for the Geology Club on campus. After graduation, Jamie plans to work a few years in industry and return to school for a graduate degree.

Jessica Sherman

Jessica Sherman is fourth year biology major with a concentration in Molecular and Biotechnology. She has been involved in ASI as the Board of Director for the College of Natural Sciences and Mathematics, the Vice President of Internal Affairs for Beta Psi Omega, and a Peer Mentor Lead for the Think Like Einstein Program. Jessica currently has an internship in the COPE Health Scholars Program and in the future she will be attending Physician Assistant School.



Judith Avila

Judith is majoring in Geology. Her research explores climate change in Northern California over the past 10,000 years. Her career goal is to work with the United States Geological Survey or become a tenured faculty at a four year university. She is in the Honors, McNair Scholars, and Louis Stokes Alliance for Minority Participation Programs. The Geology department awarded her the Outstanding Academic Achievement Award in Geology (2018) as a junior, and the Outstanding Major Award (2019) as a senior.

Karen Walsh

Karen recently received her Bachelor's degree in Evolutionary and Ecology Biological Sciences at Cal State University, Fullerton in Spring 2018. She is continuing her education at CSUF as a Master's student studying evolutionary genomics of aging in Drosophila melanogaster in Dr. Parvin Shahrestani's laboratory. Her future career goals involve teaching in higher education and becoming a Registered Veterinary Technician.

Kenneth Eaves

Kenneth completed his undergrad in mathematics with a concentration in teaching at CSUF. Now a high school math teacher, Kenneth is completing his masters in mathematics and will graduate in May. Working with Dr. Lyons, he has been analyzing the connections between weight systems that arise from geometric surfaces called K-3 surfaces and is continuing this research through May.

Kristina Rosete

Kristina Rosete is a second-year student at California State University, Fullerton. She is majoring in mathematics with a concentration in teaching. She is involved in the PRIME club and SMART Girls. Kristina worked on a research project over the summer with Dr. Soto and two other students, and presented the project at the Joint Mathematics Meeting in Baltimore where they received the Outstanding Poster Award. In the future, she hopes on becoming a high school mathematics teacher, or a professor.

Lindsey Langer

Lindsey Langer is an undergraduate geology major. She is completing an undergraduate thesis with Dr. Jeffrey R. Knott on Provenance, Detrital Zircon Age and Tephrochronology of Pliocene Conglomerate, in Eastern Deep Springs Valley, California. She is a member of Project RAISE and the geology club. She currently works as a student assistant in the Geology Department. Lindsey plans to work in consulting and obtain her Professional Geologist (P.G.) license.

Manalo Kaela

Kaela Manalo is a senior pursuing a bachelor's degree in biology, with a concentration on ecology and evolutionary biology. She is a member of Dr. Stapp's Ecology lab working with coyotes. Her future plans include attending veterinary schools, working in wildlife rehabilitation, or in a zoo.

Mark Rodriguez

Mark Rodriguez is a 4th year undergraduate student majoring in biochemistry. He is currently a full-time student who helps manage and lead a retail store part-time. His goal is to pursue a career in the medical field; with his eyes specifically set on becoming an optometrist. Mark joined Dr. Pecic's research lab in order to grasp a better understanding on the practical uses of medicinal chemistry in society.

Melanie Garcia

Melanie is majoring in biology. She is a dedicated member of Dr. Parvin Shahrestani's Evolutionary and Genomics Research (EAGR) Lab where she works towards further expanding the forefront of scientific research by working collectively with the team and collaborating on various projects with other universities using *Drosophila melanogaster*. She is also involved in expanding awareness and support to underserved communities about students pursuing an education in STEM.

Melissa Wong

Melissa Wong is an undergraduate mathematics major finishing her last semester at Cal State Fullerton. She has conducted research with Dr. Roberto Soto for the past two years. Melissa is also an executive officer of PRIME Club and SMART Girls. After graduation, she plans to attend graduate school and hopes to apply her background in mathematics to further biomedical research.

Mikayla Mays

Mikayla Mays is a graduate student in the physics department and a TA for undergraduate physics labs. She graduated from CSUF with her B.S. in Physics, and stayed on in the Master's program to continue her research in Physics Education. Once she graduates she hopes to pursue a Ph.D. and eventually a faculty position at a university.

Nayelie Benitez Santos

Nayelie Benitez Santos is a senior, graduating in the Spring of 2019 with a BS in Health Science and a minor in Aging Studies. She will be starting her MS degree in Epidemiology at UCLA in the fall of 2019. Her research interest includes using epidemiological methods to study the effects of environmental exposures on human health, such as chronic diseases.

Saul Lopez

Saul Lopez graduated from CSUF with B.A. in applied mathematics. He is currently in the mathematics master's program with emphasis in teaching. Saul was an Instructional Assistance for four years in the Math Study Hall at Santiago Canyon College. He is teaching high school mathematics at Santiago Canyon College Continuing Education program where he is also a teaching associate in the mathematics department.

Sonali Vyas

Sonali Vyas is a third-year student at Cal State Fullerton and is a mathematics major with a concentration in Modeling and Computational Applied Mathematics and English major. She has worked on two research projects with Dr. Roberto Soto and presented one at the Joint Mathematics Meeting in Baltimore. She is a board member of the PRIME club and editor of the mathematics section of Dimensions. After she receives her undergraduate degree she plans to either attend Graduate School or join the Peace Corps.

Steven Guillen

Steven Guillen is a senior majoring in biochemistry at CSUF who will be graduating this Spring, 2019. Steven's undergraduate research was completed under the guidance of Dr. Andrew Petit in collaboration with Dr. Amanda Evans. He joined the Petit Lab in 2017 and worked independently to initiate this joint theory-experimental study. Steven is excited with his plans to continue on his educational pathway in graduate school.

Suzette Herrera

Suzette Herrera is pursuing a Bachelor of Science in Biochemistry at California State University, Fullerton. She works under the guidance of Dr. Nina Robson and Dr. Madeline Rasche on the transdisciplinary team, Titanano, a collaboration between the biochemistry and mechanical engineering departments. Suzette's research includes the development of a bio-nano-robot which delivers chemotherapeutic treatments directly to cancer cells.

Taylor Salgado

Taylor Salgado is majoring in Liberal Studies with an emphasis in elementary education. She is going into her senior year and plans to further her education after. She hopes to receive her masters in education and impact the lives of her little students.

Vianney Martinez

Vianney Martinez is a Chemistry major graduating in Spring 2019. She served as the Dimensions section editor for Chemistry and Biochemistry department and as the Director for ASI's Environmental Sustainability Commission. She hopes to dedicate her time over the next year battling environmental issues, and to return to school for her masters in Environmental Studies.

Zsuzsanna Papp

Zsuzsanna Papp is a graduating Biological Science major. In addition to her work with Dr. Kristy Forsgren, Zsuzsanna is working as a laboratory assistant at SCCWPR (Southern California Coastal Water Research Project). She is a scholar in the STEER Program (Scholarships to Enhance Excellence in the Chemical and Biological Research-Based Workforce) and in the future she plans to attend medical school.