The Board of Trustees
of the
Center for Excellence in Education

CENTER FOR EXCELLENCE IN EDUCATION
8201 Greensboro Drive, Suite 215
McLean, VA 22102

Research Science Institute 2004
Faculty and Staff

RSI Directors

Mr. Matthew Paschke (MIT)	RSI '91	North Canton, OH
Dr. Carey Inouye (Caltech)		Honolulu, HI

Vice President of Administration

Ms. Maite P. Ballestero		Bethesda, MD

Assistant Directors

Mr. James Skelley (MIT)	RSI '99	Joplin, MO
Ms. Dagny Looper (Caltech)	RSI '99	Bryant, AR

Professors — MIT

Dr. Amy Bany	RSI '90E	Yale University
Mr. Adrian Campbell	RSI '92	Bayer Corporation
Dr. Leonardo Hsu	RSI '87	University of Minnesota
Mr. Lance Rhoades		University of Washington
Dr. Christopher Skinner	RSI '88	University of Michigan

Professors — Caltech

Dr. Charles Doran	RSI '87	Columbia University
Dr. Parth Patwari	RSI '90E	Brigham & Women's Hospital
Mr. Lance Rhoades		University of Washington
Dr. Karn Sorasaenee		California Institute of Technology
Dr. Dennis Ugolini	RSI '90W	Trinity College

Academic Staff — MIT

Ms. Shaudi Hosseini	RSI '99	Kingston, PA
Mr. Christopher Mihelich	RSI '97	Princeton, NJ
Ms. Viviana Risca	RSI '99	Port Washington, NY
Dr. Jenny Sendova		Sofia, Bulgaria
Mr. Balint Veto	RSI '98	Cambridge, MA

Academic Staff — Caltech

Ms. Jacqueline Choi	RSI '97	Leonia, NJ
Dr. John Rickert		Terre Haute, IN
Mr. David Zhang	RSI '99	Overland Park, KS

Teaching Assistants — MIT

Mr. Allen Chao	RSI '02	Beaumont, TX
Mr. Anatoly Preygel	RSI '02	Germantown, MD
Ms. Michelle Rengarajan	RSI '02	Pasadena, CA

Mr. Prasanna Tambe	RSI '91	Philadelphia, PA
Mr. Vivek Venkatachalam	RSI '01	Berkeley Heights, NJ
Mr. Kartik Venkatram	RSI '00	Morristown, NJ

Teaching Assistants — Caltech

Mr. Edward Boas	RSI '94	San Diego, CA
Mr. Bryant Mathews	RSI '97	Grand Rapids, MI
Ms. Emma Schmidgall	RSI '02	Golden Valley, MN
Ms. LeeAnn Tzeng	RSI '90E	Hanover, NH

Technical Assistants

Mr. Ho Seung Paul Ryu (MIT)	RSI '03	Pacific Grove, CA
Mr. David Jake Wildstrom (Caltech)	RSI '97	Kensington, MD

Student Counselors — MIT

Ms. Cynthia Chi	RSI '03	Sugar Land, TX
Mr. Logan Dean	RSI '03	Quantico, VA
Ms. Julie Finkelstein	RSI '03	Woodland Hills, CA
Ms. Aidan Parker	RSI '02	Cambridge, UK
Mr. Arup Sarma	RSI '03	Shawnee Mission, KS
Mr. Jacob Temme	RSI '02	Casper, WY

Student Counselors — Caltech

Mr. Steve Byrnes	RSI '02	Lexington, MA
Mr. Jon Hanover	RSI '03	Chesterland, OH
Ms. Corinna Zygourakis	RSI '01	Houston, TX

Preface

The *Research Science Institute 2004 Compendium* displays the work of 109 academically talented high-school students who participated in the 2004 Research Science Institute (RSI) administered by the Center for Excellence in Education with the sponsorship of the Massachusetts Institute of Technology in Cambridge, Massachusetts, and (in its inaugural year of sponsorship) the California Institute of Technology in Pasadena, California.

This unique six-week summer program combines theoretical classroom instruction in mathematics and science with one-on-one research experience with scientists from the sponsoring institutions and various nearby universities, private corporations, and research organizations in and around Cambridge and Pasadena. It is the only U.S. program of its kind offered at no cost to participants. The scientists and their respective organizations are acknowledged in the headings of the reports.

These reports are the culmination of students' research in mathematics, physics, chemistry, biology, computer science, and engineering. The faculty of the Institute, together with a team of outside experts, selected five outstanding written reports from each program site for special recognition. Panels of visiting educators and scientists also designated five projects from each site as examples of outstanding oral presentation. The chosen written reports are printed first in their entirety,[†] followed by the abstracts of the commended oral presentations, and finally by the remaining student abstracts.

The Institute has drawn students from all over the U.S. and from 45 other countries. In 2004 the 109 students represented 39 states of the U.S. and the District of Columbia, as well as the countries of Bulgaria, China, Germany, Greece, India, Israel, Jordan, Kuwait, Lebanon, Morocco, Pakistan, Poland, Saudi Arabia, Singapore, Turkey, and the United Kingdom. The students were all selected having completed coursework comparable to or exceeding that of the third year of high school. The U.S. students averaged 79 of 80 points on the mathematics section and 74 of 80 points on the verbal sections of the Preliminary Scholastic Aptitude Test (PSAT), representing the 99th percentile of all high-school students taking the PSAT.

The inestimable academic and social value of the Institute to scientifically talented students is evident from the following praises from participants of RSI 2004:

> *I have never felt better than the moment I realized I was a "Rickoid."*
>
> *RSI assured me that I am on the right path … that I will want to major in math or science and choose a career in math and science.*
>
> *I learned what real research is like.*
>
> *It made me realize that research requires a lot of patience and perseverance.*
>
> *This summer has only increased my love for the sciences.*
>
> *I had the best experience in my life; I have improved my leadership skills and made lots of friends.*
>
> *The students are incredible!*

The editors are honored to have taken part in the Institute as both students and staff and are pleased to present to you the excellent work of their successors.

Chris Mihelich and **Jake Wildstrom**, RSI 1997, *Editors*

[†] But some appendices containing computer code or other supplementary material have been suppressed to conserve space.

Table of Contents

Synthesis of 3-amino-2-alkenoate

Albert Chen

under the direction of
Dr. Mohammad Movassaghi and Robert Sindelar
Massachusetts Institute of Technology

Abstract

The main aim of the project is to synthesize an isomerically pure sample of 3-amino-2-alkenoate, a useful β-amino acid derivative that is a precursor to molecules with medicinal properties. We successfully made the product from the reaction of ethyl acetoacetate and ammonium acetate dissolved in methanol. We also attempted to synthesize 3-pyrrole-2-alkenoate from the reaction of our product and 2,5-dimethoxytetrahydrofuran. The products of both reactions were chracterized by means of ^{1}H nuclear magnetic resonance spectroscopy.

1 Introduction

Organic chemistry research has revealed the promising applications of β-amino acids and their derivatives, encouraging further study of their properties and methods of synthesis. β-amino acids rarely occur in nature, but they do appear in free form or in active peptides [9]. β-amino acids have one additional carbon in the parent chain compared to the more common α-amino acids used to make most polypeptides. Therefore, they have different properties and can be used to create new classes of organic molecules. For example, peptides composed of β-amino acids are resistant to decomposition by enzymatic activity [6], whereas α-peptides are not. There is also evidence that β-peptides may fold in similar ways to α-peptides under certain conditions, specifically when nearby monomers have weak intermolecular forces between them [5]. They have also been shown to form secondary structures similar to those of α-peptides [10]. Since β-peptides may be useful to the human body but are not broken down by digestion, they are potentially a novel class of medicinal molecules. In addition to peptides, β-amino acid derivatives can form other useful molecules. For example, β-amino-α-hydroxy acids can be used to produce taxol, a treatment for tumors; bestatin, an immune system modifier; and oligopeptides that can be used as sedatives [3]. Also, β-amino acids are used to produce β-lactams, a class of antibiotics that includes penicillin [1].

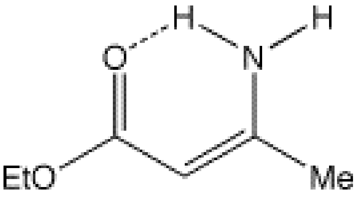


Figure 1: 3-amino-2-alkenoate

In this paper we explore an approach to the synthesis of 3-amino-2-alkenoate, a useful β-amino acid derivative (Figure 1). Using ethyl acetoacetate, a β-keto ester, as starting material, we seek a pure sample containing only the shown isomer. We also present a reaction mechanism that is plausible given the chemistry of the reaction. We also share our work on a method for the formation of the vinyl pyrrole of the 3-amino-2-alkenoate.

Thin-layer chromatography (TLC) and gas chromatography (GC) were used to track the progress of the reaction and verify that it had gone to completion. Flash chromatography (column chromatography) was used to purify the sample and to separate out components. Nuclear magnetic resonance (NMR) spectroscopy was conducted to ensure the fidelity of the product. ^{1}H NMR signals vary based on the chemical environments of protons in the molecule and therefore can be used to infer a structure.

Goals of synthesis include removal of impurities, optimization of yield, separation of isomers, and verification of product identity. β-amino acid derivative samples with isomeric purity are important in forming arrangement-dependent molecules because alternative isomers may not react identically [6]. Here, we are concerned with the cis-trans isomeric disparity that occurs with double bonds. We seek the cis isomer, with the amino group interacting with the carbonyl group. Cis-trans isomers often either are in equilibrium with each other or are produced coexisting in the same sample. One efficient method for isomeric isolation is to determine a convenient step of the synthesis process where the isomers can be separated via chromatography, because the isomers of some molecules have greater differences in physical properties than those of others.

2 Experimental Section

2.1 General Procedures

All reactions were performed in flame-dried round bottom flasks fitted with rubber septa under a positive pressure of argon. Air- and moisture-sensitive liquids and solutions were transferred via sterile syringe. Analytical thin-layer chromatography was performed using glass plates precoated with 0.25 mm 230–400 mesh silica gel impregnated with a fluorescent indicator (254 nm). Thin-layer chromatography plates were visualized by exposure to ultraviolet light and to an indicator solution of *p*-anisaldehyde followed by heating on a hot plate. Organic solutions were concentrated by rotary evaporation at reduced pressures and high temperatures.

2.2 Instrumentation

Proton NMR spectra were recorded with a Mercury Varian (300 MHz or 500 MHz). Chemical shifts for protons are reported in parts per million downfield from tetramethylsilane and are referenced to residual protium in deuterated chloroform or benzene. NMR data are organized as follows: chemical shift, multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet), integration, coupling constants in Hertz, and assignment.

2.3 Materials

Ethyl acetoacetate was purchased from Aldrich and distilled within the laboratory. The methanol and ammonium acetate were purchased from Fisher and Alfa Aesar respectively. 2,5-dimethoxytetrahydrofuran, 10-camphorsulfonic acid, and all organic solvents were purchased from Aldrich.

Figure 2: Ethyl acetoacetate and 3-amino-2-alkenoate

2.4 3-amino-2-alkenoate

The reaction to produce the β-amino acid derivative (Figure 2) was conducted as follows. A 250 mL round-bottom flask (RBF) with a magnetic stir bar was flame-dried using a propane torch. Methanol (150. mL) was poured into the RBF along with solid ammonium acetate (46.2 g). The flask was then covered with a septum and evacuated and flushed with argon three times and then put under positive argon pressure. 15.3 mL of ethyl acetoacetate was syringed into the flask. The flask was placed over a magnetic spinner and stirred. Theoretical yield was calculated to be 15.5 grams.

Within an hour after the reaction had been begun, a 0.1 mL sample was taken from the flask and diluted with dichloromethane in order to execute a TLC (solvent: 25% EtOAc/hexanes) and a GC (maximum retention time: 15 minutes). Both demonstrated that the reaction was proceeding rapidly. The reaction was continued for another 15 hours before another sample determined that the reaction was complete.

The flask was then placed on a rotary evaporator to evaporate excess solvent. Water was observed on the edges of the flask. The product was then diluted with 300. mL of dichloromethane, which produced a solid white precipitate. The precipitate was vacuum filtered and then rinsed with 2 × 300. mL of dichloromethane. The filtrate was poured into a separatory funnel and washed first with distilled water and then with brine. Because dichloromethane is denser than water, the organic layer and therefore the product sank to the bottom. The organic layer was collected in an Erlenmeyer flask and dried over anhydrous sodium sulfate. The product was then vacuum-filtered again and then concentrated in a 1000 mL RBF using a rotary evaporator. After the solvent had evaporated, the yield of the crude 3-amino-2-alkenoate was 15.0 grams.

A 5 cm-wide chromatography column was prepared as described by McMurry [8]. Approximately 1500 mL of solvent was made from 10% EtOAc/hexanes. Anhydrous sodium sulfate was poured into the bottom followed by 8 cm of compressed silica gel. After more sodium sulfate was added to the top, the product was loaded onto the column and was allowed to submerge into the silica before the rest of the column was filled with solvent. 30 mL fractions were taken and a TLC was done on each. Fractions containing the material that matched the characteristics of the product were pooled into a preweighed RBF. Fractions containing other material were pooled separately. The RBF was placed on a rotary evaporator to obtain the product. The final yield of the pure 3-amino-2-alkenoate was 12.0 grams (77.4%).

NMR spectroscopies were taken of the 3-amino-2-alkenoate. The placement of the hydrogens in the molecule showed that both the cis and trans forms were present, most likely in rapid equilibrium with each other. Because of the fast cis-trans isomerization of 3-amino-2-alkenoate, it was unlikely that the isomers would be separable. Therefore, the combined isomers were reacted to synthesize 3-pyrrole-2-alkenoate considering its slower isomerization.

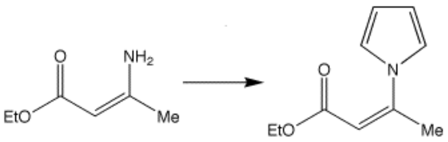


Figure 3: 3-amino-2-alkenoate and 3-pyrrole-2-alkenoate

2.5 The Vinyl Pyrrole

The synthesis of 3-pyrrole-2-alkenoate (Figure 3) was attempted. 500.2 mg of the 3-amino-2-alkenoate product of the previous reaction was weighed out and placed in a 100 mL round-bottom flask. 90.02 mg of solid 10-camphorsulfonic acid was added, and then the flask was covered with a septum. It was subsequently placed under vacuum and flushed with argon three times before being put under positive argon pressure. 38.7 mL of THF was syringed into the flask in order to make the molarity of 3-amino-2-alkenoate 0.1 M. Then 0.500 mL of 2,5-dimethoxytetrahydrofuran was syringed to start the reaction. Theoretical yield was calculated to be 694 mg.

Within one hour, the reaction mixture was sampled for TLC and GC. Both indicated the reaction was complete. Two main compounds seem to have been formed. We hypothesized that they were the cis and trans isomers of the vinyl pyrrole. The molecules stained and separated in the same way as the TLC standard, but the retention times were drastically different from the GC standard for the vinyl pyrrole.

The reaction mixture was worked up by washing with 25 mL water and 50 mL ethyl acetate and separation in a separatory funnel. Any extra product remaining in the aqueous layer was extracted by washing with 3 × 25 mL ethyl acetate in the separatory funnel. All organic layers were then combined and dried over sodium sulfate. The product was then filtered by vacuum to remove sodium sulfate and other solid impurities. Solvent was then evaporated under reduced pressure.

Flash chromatography was performed on the product using a 3 cm-wide column and 10 mL fractions. The solvent system was 10% EtOAc/hexanes. One pool provided one of the major products, another provided a product not previously seen and that appeared to be starting material, and no pool provided the second major product. Impure fractions were pooled together to form a final pool, which was subjected to a second flash chromatography after being concentrated. The second column utilized 3% EtOAc/hexanes as solvent in order to induce greater separation. 482 mg of product was recovered.

NMR spectroscopy indicated that neither of the primary products of the reaction were the correct product, implying that unwanted hydrolysis had occurred. A second reaction, using acetic acid instead of 10-camphorsulfonic acid, gave the same GC results; presumably the same compounds were formed.

3 Results

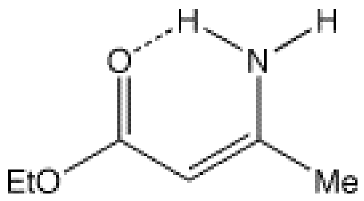
3.1 3-amino-2-alkenoate

The ethyl acetoacetate and the product were both UV-active and had virtually the same Rf (retention factor) values. In 25% EtOAc/hexanes, the Rf value was 0.31. Anisaldehyde caused the starting material to stain white surrounded by red. The product stained red. These agreed with chromatographic standards, so TLC indicated that the product was

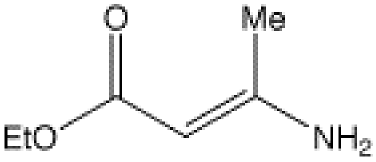
3-amino-2-alkenoate.

The retention times of ethyl acetoacetate were 1.624 and 2.269 minutes. Ethyl acetoacetate does not behave well in GC procedures, changing between the enol and ketone forms and creating two retention times. The retention time of the product was 4.330 minutes. This agreed with chromatographic standards, so GC indicated that the correct product had been formed.

NMR results are as follows.



^{1}H NMR (500 MHz, $CDCl_3$, 20 °C): 7.93 (bs, 1H, N**H**), 4.54 (bs, 1H, N**H**), 4.52 (s, 1H, C=C**H**), 4.11 (q, 2H, *J*= 7 Hz, CH_3C$\mathbf{H_2}$), 1.91 (s, 3H, C$\mathbf{H_3}$C=C), 1.26 (t, 3H, J= 7 Hz, C$\mathbf{H_3}$$CH_2$).



^{1}H NMR (500 MHz, C_6D_6, 20 °C): 5.00 (s, 1H, C=C**H**), 4.21 (q, 2H, *J*= 7 Hz, CH_3C$\mathbf{H_2}$), 2.27 (s, 3H, C$\mathbf{H_3}$C=C), 1.28 (t, 3H, *J*= 7 Hz, C$\mathbf{H_3}$$CH_2$), NH_2 protons not observed.

NMR showed that both cis-trans isomers were present in the sample. This was inferred from the two different peaks for the protons in the methyl group and two different peaks for the protons in the amino group. The hydrogen bonding of the amino group to the nearby carbonyl group produced the noticeable difference.

3.2 3-pyrrole-2-alkenoate

The products of the attempted synthesis of the vinyl pyrrole were not UV-active and had Rf values of 0.31 and 0.41, respectively. These agreed with chromatographic standards for the two cis-trans isomers of 3-pyrrole-2-alkenoate, indicating that the correct products had been formed.

The retention times of the products of the attempted synthesis of 3-pyrrole-2-alkenoate had retention times of 2.307 and 2.451 minutes, respectively. These disagreed with the chromatographic standards of approximately 5 minutes. These products were much more volatile than the desired products, demonstrating that the incorrect products had been formed.

NMR results showed that the incorrect products had formed.

4 The Mechanism

A reaction mechanism for the formation of 3-amino-2-alkenoate was proposed (Figure 4). Ammonium acetate is in equilibrium with ammonia and acetic acid. Even though the salt is favored in the equilibrium, the ammonia and acetic acid can still react in the mechanism of the formation of 3-amino-2-alkenoate. The acetic acid protonates **1** because of the high density of electrons around the electronegative oxygen. The extra proton further increases the tendency for electrons to be drawn towards the oxygen, making the carbon nearest to it electron-deficient. This gives the nucleophilic ammonia the ability to donate electrons to the carbon to form **3**. The acetate ion, attracted by charge to the amino group, then pulls off a hydrogen to regenerate acetic acid. Once again, the acetic acid protonates the oxygen due to the oxygen's effective negative charge. Once **5** is achieved, the protonated oxygen makes up a very likely leaving group, and is released as water, taking the carbon-oxygen bond electrons with it. The lone-pair electrons on the amino group are forced to bond with the carbon to stabilize the compound, thus forming **6**. Because the amino group is positively charged, it attracts the acetate anion and gives up its proton to form **7**. In the final step, tautomerization occurs as a hydrogen is transferred from within the molecule to the amino group, causing a rearrangement of bonds. Because of the stability of the hexagonal ring, 3-amino-2-alkenoate is unlikely to participate in the reverse reaction.

Overall, the acetic acid is not consumed and is therefore a catalyst. The ammonia, however, reacts with the ethyl acetoacetate. Water is formed as a product, which was observed in the reaction. The reaction does not go to completion quickly; it takes several hours to finish, which is perhaps due to the reversibility of each step of the mechanism. The mechanism fits the observations we made.

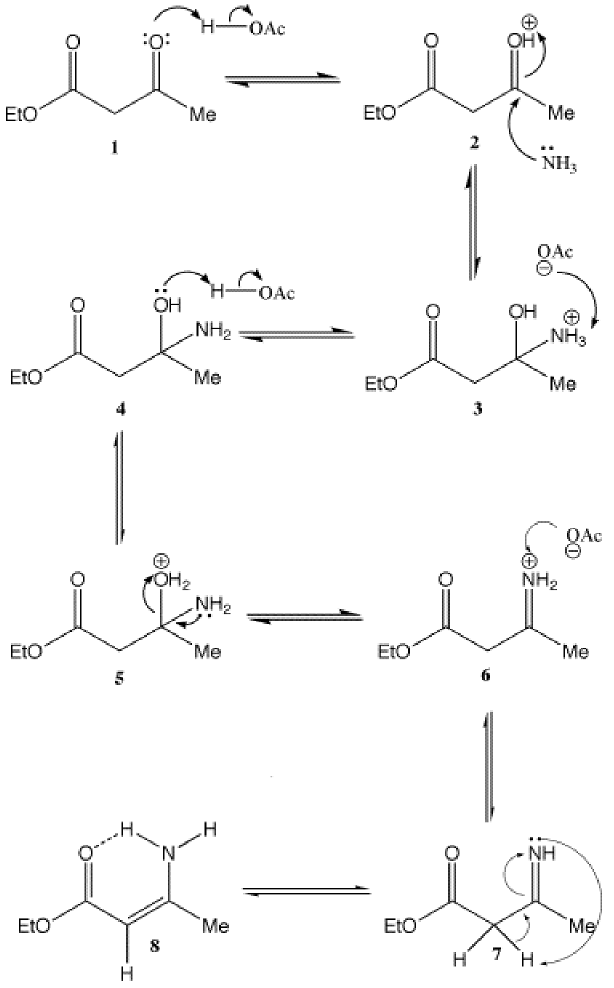


Figure 4: Proposed mechanism for the production of 3-amino-2-alkenoate

5 Future Studies

Further investigation into the viability of the method of synthesizing 3-pyrrole-2-alkenoate from 3-amino-2-alkenoate is needed. Research into the production of β-amino acids from the molecules synthesized here could provide this work with numerous medicinal applications. Other pathways from ethyl acetoacetate to 3-pyrrole-2-alkenoate should also be studied (Figure 5).

6 Conclusion

We have found that the synthesis of 3-amino-2-alkenoate from ethyl acetoacetate is a viable and moderately efficient method and presented a mechanism explaning the reaction's success. This product was purified and has been isolated with 77.4% yield. We have also found, however, that one method of producing 3-pyrrole-2-alkenoate from 3-amino-2-alkenoate is not viable. Our findings build onto the growing knowledge pool of how to best synthesize β-amino acid derivatives for eventual commercial or industrial use.

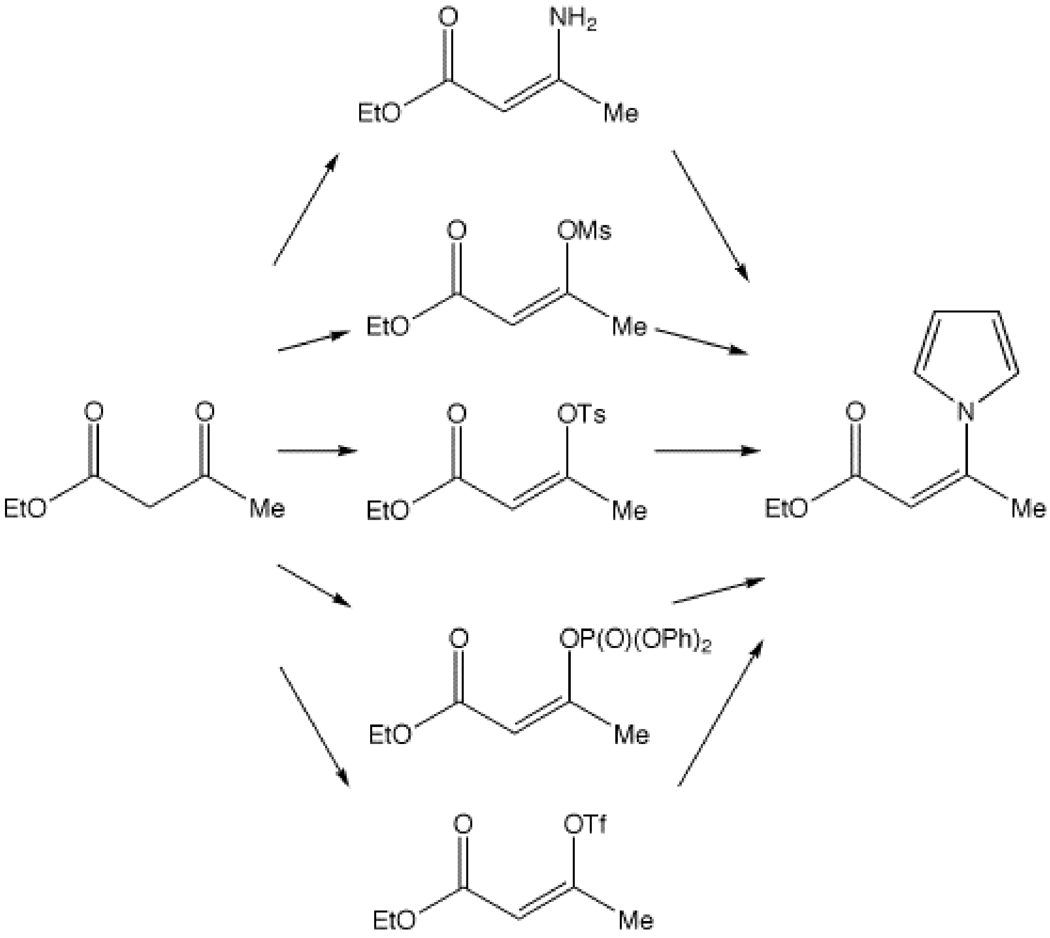


Figure 5: Other pathways currently being explored

7 Acknowledgments

I would like to thank my mentor Professor Mohammad Movassaghi for teaching me the concepts of organic chemistry and overseeing the direction of my project. I would also like to thank my laboratory mentor Robert Sindelar for demonstrating, teaching, and supervising the lab techniques of my research and for getting me the materials and information I needed. My fellow lab workers deserve credit as well, especially Allison, Matt, Giovanni, Diana, and Mike.

I thank Chris Mihelich for editing my paper and giving me advice on my mentorship and my presentation. I acknowledge Paul Ryu for helping me with technical and computing issues. I acknowledge Alan Chao and Kartik Venkatram for teaching me how to use Athena and LaTeX to write my paper.

Finally, I would like to thank the Research Science Institute and the Center for Excellence in Education for providing me with the opportunity to perform this research and immersing me in a stimulating intellectual environment.

References

[1] C. Gomersall. Beta-lactams. Available at `http://www.aic.cuhk.edu.hk/web8/beta-lactams.htm` (2004/07/23).

[2] D.L. Pavia, G.M. Lampman, and G.S. Kriz. *Organic Laboratory Techniques.* 3rd ed. Saunders College Publishing, Philadelphia, PA (1998).

[3] G. Cardillo and C. Tomasini. Asymmetric synthesis of β-amino acids and α-substituted β-amino acids. Chemical Society Reviews (1996), 117–128.

[4] G.M. Loudon. *Organic Chemistry.* 3rd ed. The Benjamin/Cummings Publishing Company, Inc., Redwood, CA (1995).

[5] G.P. Dado and S.H. Gellman. Intramolecular hydrogen bonding in derivatives of β-alanine and γ-amino butyric acid: Model studies for the folding of unnatural polypeptide backbones. Journal of the American Chemical Society 116 (1994), 1054–1062.

[6] G. Zhu, A. Chen, and X. Zhang. Highly efficient asymmetric synthesis of β-amino acid derivatives via rhodium-catalyzed hydrogenation of β-(Acylamino)acrylates. Journal of Organic Chemistry 64 (1999), 6907–6910.

[7] J. Kling. Beta-peptide antimicrobials. Available at `http://www.chemistry.org/portal/a/c/s/1/feature_pro.html?id=5ef69da2d21111d5e8e54fd8fe800100` (2004/07/03).

[8] J. McMurry. *Organic Chemistry.* 4th ed. Brookes/Cole Publishing Company, Pacific Grove, CA (1996).

[9] N. Sewald. Synthetic routes towards enantiomerically pure β-amino acids. Angewandte Chemie International Edition 42 (2003), 5794–5795.

[10] S. Krauthauser, L.A. Christianson, D.R. Powell, and S.H. Gellman. Antiparallel sheet formation in β-peptide foldamers: effects of β-amino acid substitution on conformational preference. Journal of the American Chemical Society 119 (1997), 11719–11720.

Regulation of Late-Stage Flower Development by Downstream Genes of the Homeotic Protein AGAMOUS

Jerry Guo

under the direction of
Dr. Toshiro Ito and Professor Elliot Meyerowitz
Division of Biology
California Institute of Technology

Abstract

The genes and pathways AG regulates in late-stage flower development are largely unknown. Several putative downstream genes of AG involved in anther dehiscence were identified, including DAD1, MYB26, OPR3, COI1, and RBR1, through a bioinformatics approach. A 35S::AG-GR inducible line was constructed for timed-induction expression analysis of anther dehiscence versus indehiscence, which revealed the likely direct induction of DAD1 by AG. Mutated AG binding sites of DAD1 and MYB26 were linked to the β-glucuronidase reporter gene; the plasmid constructs were transformed into plants for GUS staining to test *in vivo* site activity. Results support the hypothesis that AG is continually functional and controls late-stage flower development by regulation of downstream genes.

1 Introduction

Floral organ development is a phenomenon that spatially and temporally involves the specific expression of a considerable number of genes in a variety of cascades and networks. Since the rise of the molecular age, much basic research has been devoted to elucidating the mechanisms and functions of and relationships between these homeotic genes [1]. Fortunately, because the evolutionary process provides relative physiological and biochemical similarity in plants, fundamental discoveries on the model organism *Arabidopsis thaliana* can be widely applied to all 25,000 flowering species. *Arabidopsis* not only produces a large number of seeds in a short generational period, but its entire genome has also been sequenced and mutant lines can be established [2].

1.1 Floral Organ Structure

Angiosperm eudicots develop flowers to produce gametophytes. The flower consists of four concentric whorls of organs: (1) sepals, (2) petals, (3) stamens, and (4) carpels, that sequentially form inward from the floral meristem. Stamens, the male reproductive organ, and carpels, the female counterpart that holds the ovules, produce sporogenous cells around floral stage 5. The stamen consists of an anther, the site of pollen development, and a stalk-like filament that aids in pollen dispersal. At floral stage 14, the flower opens and anther dehisces to release the mature pollen [3].

1.2 Developmental Model

The unifying principle in the flower-development field is the ABC model, which is applicable to a wide range of species. Three classes of homeotic selector genes that encode transcription factors (termed A, B, and C) determine floral organ identity combinatorially. Class-A genes alone specify sepals, A and B together specify petals, B and C together specify stamens, and C alone specifies carpels. In addition, classes A and C are mutually repressive. In general, gene products are expressed in floral regions that exhibit defects in mutants. The resulting transcription factors regulate floral development through a hierarchy of downstream targets [4].

Figure 1: *Arabidopsis* plants ready for tissue sampling or mutant selection [5]

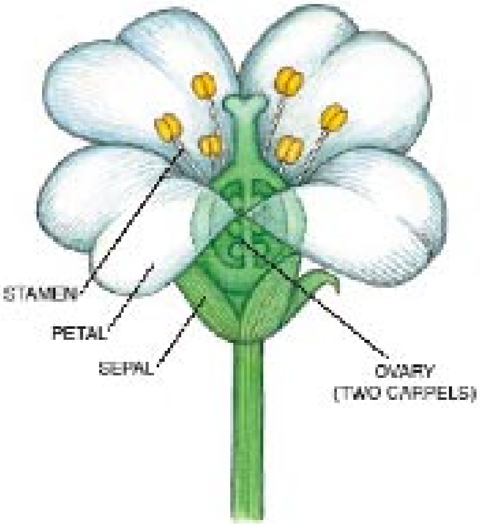


Figure 2: Four components of a flower [5]

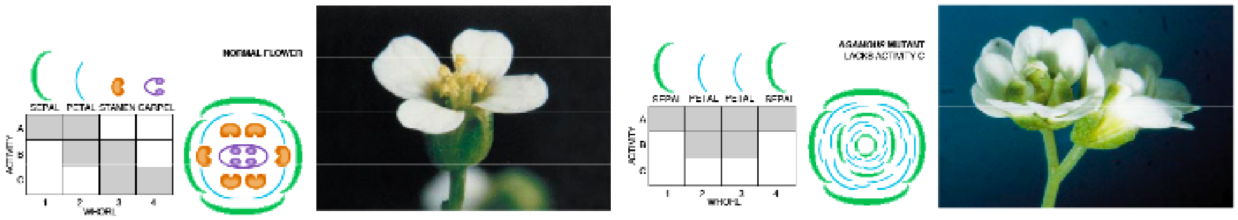


Figure 3: Flower structure of wild-type and AG mutant, as explained by the ABC model [5]

AGAMOUS, a class-C gene that encodes a transcription factor of the MADS-box family, plays a central role in flower development. The MADS-box family transcription factors bind to a target DNA sequence, the CArG box, as dimers. Since AG is responsible for reproductive organ development, mutants have a whorl pattern of sepal-petal-petal-sepal. The processes, mechanisms, and genes involved in differentiating whorl 3 and 4 primordia by AG into stamens and carpels are largely unknown [5].

1.3 Late-stage Regulatory Genes in Stamen Development

Like any other developmental process, a hierarchy of genes regulates anther dehiscence. During this late stage, programmed cell death ruptures the stomium; the tapetum and middle layer degenerates; the endothecial layer expands, and connective cells form fibrous bands. A lipid-derived signaling compound found in many plant species, jasmonic acid, was recently discovered to be actively involved in the control of anther dehiscence and male gametophyte development [3].

Three genes directly regulate this pathway and hence late-stage floral development. The *DEFECTIVE IN ANTHER DEHISCENCE1 (DAD1)* gene catalyzes the first step of JA biosynthesis by encoding a phospholipase A1 [6]. *OPR3* encodes an isozyme of 12-oxophytodienoate reductase that reduces OPDA for JA biosynthesis [7]. Mutants of both genes are sterile but can be rescued by exogenous JA. The JA signal transduction gene *COI1* encodes an F-box protein insensitive to JA treatment [8,9].

On the other hand, *MYB26* controls the mechanical rather than the biochemical aspect of anther dehiscence. Mutants produce fertile pollen when mechanically released from the anthers. Therefore, not only does the gene directly regulate late-stage development, it is also a valuable tool for manipulating male sterility. However, the endothecial layer lacks wall fortifications and does not shrink, which is necessary for retraction of the anther wall. The gene encodes a putative R2R3-type MYB transcription factor that activates the phenylpropanoid pathway to provide lignin residues for wall thickening [10].

1.4 Purpose and Significance

The genes and developmental processes that AG regulates throughout reproductive-organ development are in large part still left to be discovered. A key question to be answered is: does AG, which represents a typical plant homeotic protein, act as a switch that activates various developmental processes which then proceed without its expression, or does AG continually need to be expressed? In animals, the latter tends to be true, as illustrated by the *Drosophilia Ultrabithorax (UBX)* homeotic gene [11].

The question leads to two extreme hypotheses on how AG functions. One is that AG triggers a few genes at the top of a cascade during floral development, and in turn these AG targets regulate other genes. However, other than SHATTERPROOF2, no putative target genes of AG are known. The second possibility that is explored in this study is that AG directly and continually regulates the expression of many genes with different functions that act throughout reproductive-organ development.

Since AG has been recently implicated in microsporogenesis, an early-stage developmental process [12], it is essential to determine the expression and regulation of downstream AG genes in late-stage development in order to validate the second hypothesis. Therefore, this research will identify and explore the relationships of downstream genes of AG that regulate anther dehiscence, a late-stage developmental process. In particular, this study will determine whether the downstream genes are direct or indirect targets of AG.

AG has homologs in a variety of plant species, many of which are of great economic interest, and also in humans as oncogenes involved in the onset of cancer. Because 80% of plant food consists of flower components and only 50% of the genetic potential has been achieved, the research holds significant applications to improving food yield, genetically engineered crops, and breeding programs. At the basic-science level, the findings will also help elucidate the general behavior of plant homeotic proteins and further the floral developmental model [2].

2 Materials and Methods

2.1 AG-Inducible System

To reveal downstream processes of AG, a strain with AG-inducible activity was constructed. This strain was homozygous for the *ag-1*-null mutation and transgenic for 35S::AG-GR, a promoter-driven constitutive AG gene with a fusion at the carboxy-terminal to the steroid-binding domain of the rat glucocorticoid receptor. This genotype of *ag-1* mutants results in flowers with an indefinite sepal-petal-petal whorl pattern [13].

However, continual treatment with dexamethasone (DEX), a steroid hormone, leads to a translocation from the cytoplasm to the nucleus of the fusion protein. The resulting flower has functional stamens and carpels and resembles the *35S::AG* plants, which implies that the AG-GR fusion protein can function as a normal AG protein. On the other hand, a single treatment with DEX only partially rescues the phenotype as flowers contain petals with stamenoid structures.

2.2 Plant Materials and Treatments

All plants were grown at 22°C under constant illumination in the Landsberg *erecta* background. Transgenic plants were generated by *Agrobacterium*-mediated infiltration. The 35S::AG-GR construct was transformed into wild-type plants and the primary T1 transformants were screened by basta selection. Phenotypic analyses in the *ag-1* were performed in the T1 generation, as well as in the T2 generation that was produced by selfing *ag-1*/+ T1 plants.

2.3 AG Downstream Genes and Binding Site Identification

A review of previous research identified several genes directly implicated in anther dehiscence. To determine if these genes are downstream of AG requires the identification of putative AG binding sites to the gene

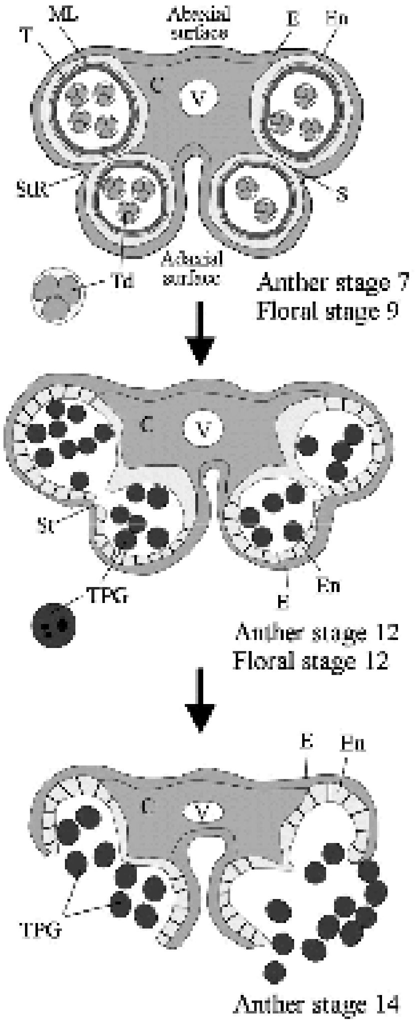


Figure 4: Cross-section of anthers at different stages. [1]C, connective; E, epidermis; En, endothecium; ML, middle layer; S, septum; St, stomium; StR, stomium region; T, tapetum; Td, tetrads; TPG, tricellular pollen grains; V, vascular bundle [3]

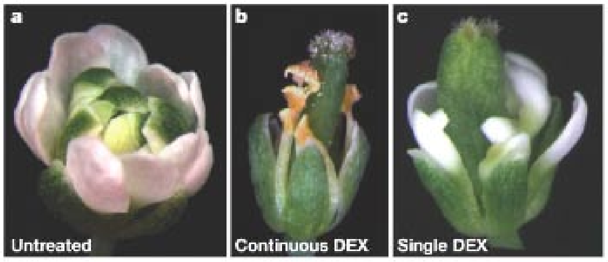


Figure 5: Phenotypic effects of *35S::AG-GR* induction

promoter. The genomic region of each identified gene was examined for the 16–base-pair perfect and weak consensus binding sequences of AG, which contains the 10-bp 'CArG-box core' (5'-CCNN(A/T)4GG-3') [14].

2.4 Timed-Induction Gene Expression Assay

Ag-1 35S::AG-GR plants were DEX treated (10 μM DEX, 0.015% Silwet L-77) either four times to induce AG and anther dehiscence or three times to observe indehiscence, in one-day intervals starting at stage 3. Inflorescence containing floral buds were collected from stage 8 to stage 13. RNA was isolated and purified from tissue samples of two independent sets by the RNeasy Mini Kit (Qiagen). Total RNA was reverse-transcribed to synthesize cDNA for RT-PCR by the ThermoScript RT-PCR system (Invitrogen).

The induction of the identified genes DAD1, MYB26, OPR3, COI1, RBR1, and a lipase control was tested through transcriptional expression levels in semiquantitative polymerase chain reaction with reverse transcription (RT-PCR) using primers DAD1-5', DAD1-3', MYB26-5', MYB26-3', OPR3-5', OPR3-3', COI1-5', COI1-3', RBR1-5', RBR-4', LIPASE-5', and LIPASE-3' after 42, 26, 28, 28, 28, and 24 cycles, respectively. Electrophoresis gel was run on the PCR products to measure the level of gene expression.

2.5 35S::AG-GR Construct

The coding region of AG was amplified from inflorescence cDNA and cloned into a blunt-ended EcoRI site of pBluescript SK (Stratagene) to produce pSK-AG. To produce pSK-AG-GR, pSK-AG was digested with EcoRI, filled in, and ligated with a DNA fragment containing a rat glucocorticoid hormone binding domain, which was excised with BamHI and XbaI from pRS020 (pDeltaGRBX) and filled in. The AG-Gr fragment was released from pSK-AG-GR by digestion with XbaI and ClaI and ligated into the corresponding sites of the binary vector pMAT137 that contains tandem cauliflower mosaic virus 35S enhancers and a terminator.

2.6 Plasmid Reporter-Gene Construct and Mutagenesis

The GUS-DAD1-5' construct was produced by first amplifying the coding region of the 3.5 kb genomic

AGAMOUS

Transcription Factors

SPOROCYTELESS/NOZZLE Regulatory Genes Regulator Genes Regulatory cascade?

Microsporogenesis Filament Elongation Anther Morphogenesis Anther Dehiscence

Figure 6: The "switch" hypothesis with late-stage functions unknown

sequence comprising DAD1 and the flanking intergenic sequences from inflorescence cDNA using primers DAD1-5'A20 and DAD1-5'B. UltraPfu High-Fidelity DNA polymerase (Stratagene) and wild-type Columbia genomic DNA were used through PCR at an extension time of 4 minutes and then cloned into the pENTR vector and subsequently into pBG-WFS7 for basta selection. An additional construct was further mutagenized by reversing the CArG-box core sequence `CC...GG` with the primers DAD1-5'MA and DAD1-5'MB and corrected to the perfect consensus with DAD1-5MCA and DAD1-5MCB using the QuikChange II XL Site-Directed Mutagenesis Kit (Stratagene).

The MYB26 promoter cloning was accomplished with amplification of the coding region by primers MYB26-5A2 and MYB26-5B at an extension time of 7 minutes. The fragment was cloned into pENTR, digested with Bst98I, blunt-ended with T4 DNA polymerase, and the GUS fragment was cut out with NarI, then blunt-ended with Klenow fragment, and finally ligated to produce the MYB26-GUS fusion. Further mutagenized and mutagenized to perfect AG-consensus sequence constructs were made by MYB26 mutagenesis primers. The MYB26-GUS fragment was then cloned into the pKGW gateway vector for plant transformation and selection. All plasmid constructs were sequenced using DAD1 primers 1–7 and MYB26 primers 1–19.

2.7 Western Blot

To check the timing of AG function, dynamic accumulation of AG-GR protein in the nuclei caused by repeated DEX treatments was examined. The *ag-1* 35S::AG-GR transgenic plants were treated once or 4 times with DEX, and the inflorescences were harvested at stage 1-10 after 3, 5, 7, 9, and 12 days. The crude nuclear extraction and purification were performed according to previously published protocols (Ito, 1997). About 50 μg of crude nuclear extra was loaded onto the 7.5% Ready polyacrylamide gel (Biorad) and subsequently blotted onto the PVDF nylon membrane (Biograd). The AG-GR fusion protein was detected with AG-specific antibody using SuperSignal west dura extended duration substrate (Pierce). The band intensity was normalized with band intensities of small nuclear proteins by Memcode reversible protein stain kit for PVDF membranes (Pierce).

2.8 GUS Staining

Stage 8 to 13 inflorescences were rinsed and stained with x-gluc to determine β-glucuronidase activity for GUS expression analysis. Chlorophyll was removed by an ethanol series. Cleared whole-mount observation was performed according to published protocols.

3 Results

Figure 7 indicates that near-consensus AG binding sites were found on several putative downstream genes. In particular, the conserved core CArG-box sequence in the 5' promoter region of DAD1 is very close to the coding region, which makes the gene a likely candidate for AG binding. MYB26 also is a likely candidate since it has two near-consensus AG binding sites on each promoter region.

The idealized Graph 1 was based on the RT-PCR products shown in Figure 8, where a lipase derivative gene was the control. Four DEX treatments induced anther dehiscence and caused a peak in DAD1 expression around day 8, which thereafter gradually declined. Three DEX treatments resulted in a much lower expression level in DAD1, which quickly disappeared. MYB26 expression, independent of AG induction, reached a sharp peak around day 6 but also

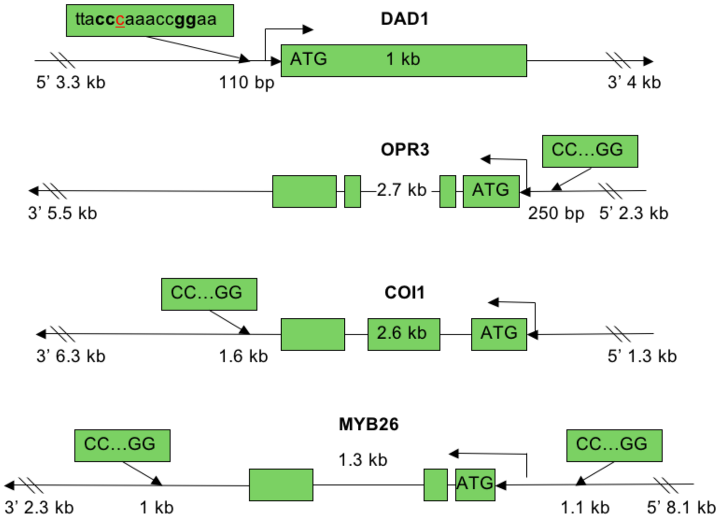
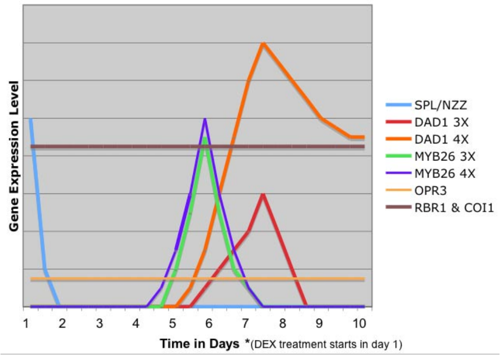


Figure 7: Sequence Schematics for Identified Putative AG Downstream Genes



Graph 1. Timed-Induction Gene Expression

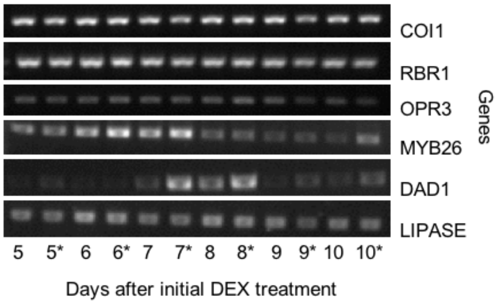


Figure 8: RT-PCR Expression Products. The asterisk denotes four DEX treatments, which ultimately cause anther dehiscence.

quickly disappeared. OPR3, RBR1, and COI showed no sign of change in expression level.

Figure 9 traces normal stamen development with prolonged AG induction. At the molecular level, four DEX treatments transferred enough AG-GR fusion protein into the nucleus where it stayed to cause anther dehiscence, as shown by Figure 10. Figure 11 focuses on the stage of anther dehiscence and reveals that the threshold for successful completion of development is four DEX treatments. Figure 12 illustrates the correct progression of flower development with AG induction.

4 Discussion

4.1 Timed-Induction Expression of Identified Regulatory Genes

The primary results, illustrated in Graph 1, of the AG timed-induction gene-expression assay reveals a substantial amount of interesting information about the relationship and functions of genes in the late-stage developmental pathway. This component of the study seeks to elucidate the effects of AG in late-stage development through an inducible system by examining the expression levels of various putative downstream genes when AG is turned on and off. First, when AG is induced, the SPOROCYTELESS/NOZZLE (SPL/NZZ) gene, which is responsible for microsporogenesis, is quickly expressed. The early, rapid induction and subsequent dramatic decline in SPL/NZZ expression implicates AG in early-stage development.

The remaining genes all regulate anther dehiscence, a late-stage developmental pathway. A pronounced difference in expression levels of DAD1 is evident between AG-induced anther dehiscence and indehiscence. The difference can be attributed to the three versus four DEX treatments. In other words, AG induction is responsible for the subsequent induction of DAD1, which supports the hypothesis of prolonged AG regulation in flower development. When AG induction is not strong enough to regulate anther dehiscence, downstream expression of DAD1 reaches a small peak shortly before the stage when anther dehiscence occurs and then quickly disappears. However, when AG does induce anther dehiscence through regulation of DAD1, DAD1 expression levels climb to a much higher peak and only drop slightly. The continued expression of DAD1 implies the necessary prolonged induction of AG for anther dehiscence to occur.

On the other hand, the MYB26 gene does not respond to a change in AG induction, as indicated by very similar expression levels between three and four DEX treatments. However, at this point, the results reveal earlier expression of MYB26 than DAD1, even though both regulate the anther-dehiscence pathway. The difference between the two genes is evidence that they are involved in separate regulatory cascades. Since the RNA expression assayed inflorescences were collected in bulk across various stages of development, differences in expression levels could be overshadowed. Thus, results from the GUS staining, which reveals gene activity spatially and temporally for individual inflorescences, must be analyzed before MYB26 can be classified as a direct, indirect, or non–AG-downstream target.

Additionally, OPR3, RBR1, and COI1 do not seem to be transcriptionally regulated at all due to the uniform levels in expression, which is in line with the negative lipase control, shown in Figure 8. Since the genes are involved in intermediate steps of various developmental pathways, it is logical that they would be continually expressed unless a homeotic regulator turns off the entire pathway. The lack of AG dependency in MYB26, OPR3, RBR1, and COI1 support the working hypothesis that AG specifically regulates certain genes in late-stage development. In this case, AG is on top of the anther dehiscence regulatory hierarchy and likely controls DAD1 as a direct target.

4.2 Phenotypic Duration Analysis

Phenotypes of AG-induced inflorescences were also examined in addition to exploring late-stage developmental relationships and functions at the molecular level. Figure 9 reveals that normal reproductive-organ development, especially in the late stages, requires continued AG induction. Prolonged AG activity throughout floral development indicates that AG actively functions in and even controls late-stage pathways. Prolonged AG activity can be seen in Figure 10, which shows the continued presence of the homeotic protein AG-GR fusion in the nucleus when induced. During anther indehiscence, there is no AG activity, as the fusion proteins quickly dissipated in the first few days, which means all downstream pathways are shut down, also illustrated by the expression assay. Therefore, the phenotypic results logically justify the molecular findings that AG controls anther dehiscence by direct regulation of downstream genes.

Furthermore, Figure 11 and 12 indicate that anther dehiscence occurs between the third and fourth DEX treatments. Obviously, no DEX treatment results in no AG induction and therefore no stamen development at all, reflective of the *ag-1* mutant. The fact that more than three DEX treatments are needed for anther dehiscence is evidence that AG is deeply involved in that developmental pathway. In addition, since just one treatment difference days after initial induction determines whether development completes successfully, AG is a critical regulator that controls a cascade of other regulatory genes in late-stage development, and the absence of AG has been seen to have far-reaching consequences during those stages.

5 Conclusions

5.1 Late-Stage Developmental Hierarchy

This study revealed that AGAMOUS is indeed functional in late-stage flower development. In particular, AGAMOUS was shown to directly control anther dehiscence through regulation of the downstream gene

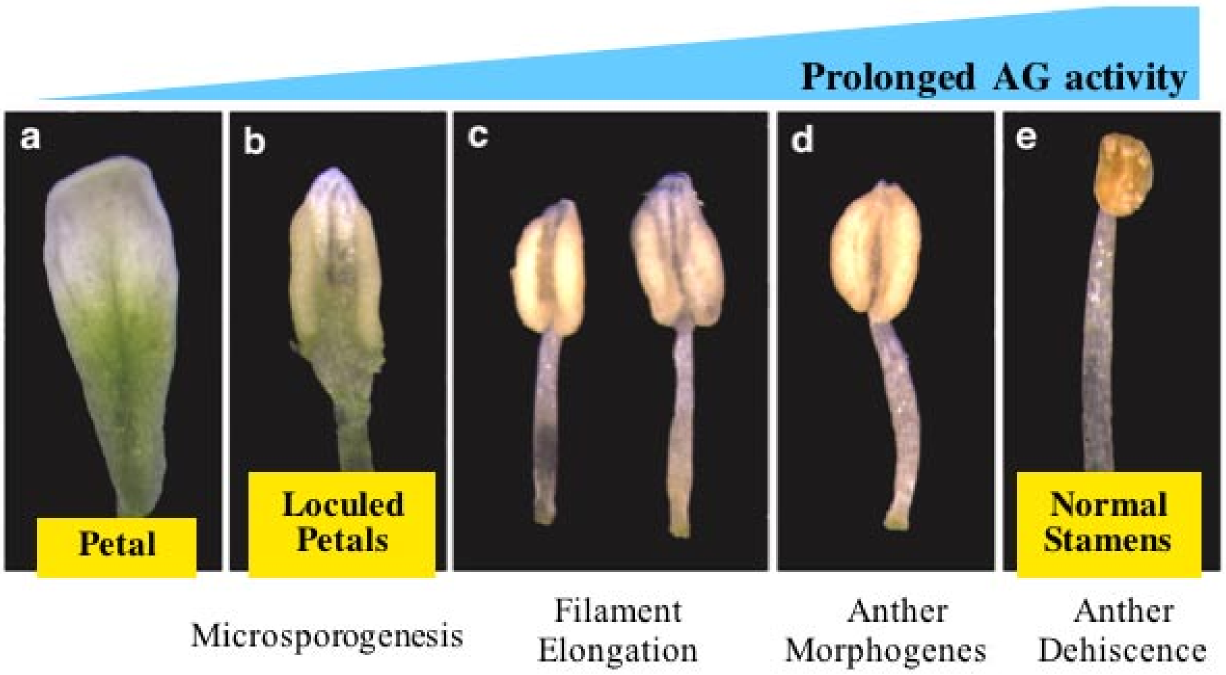


Figure 9: Continued AG Function Throughout Development

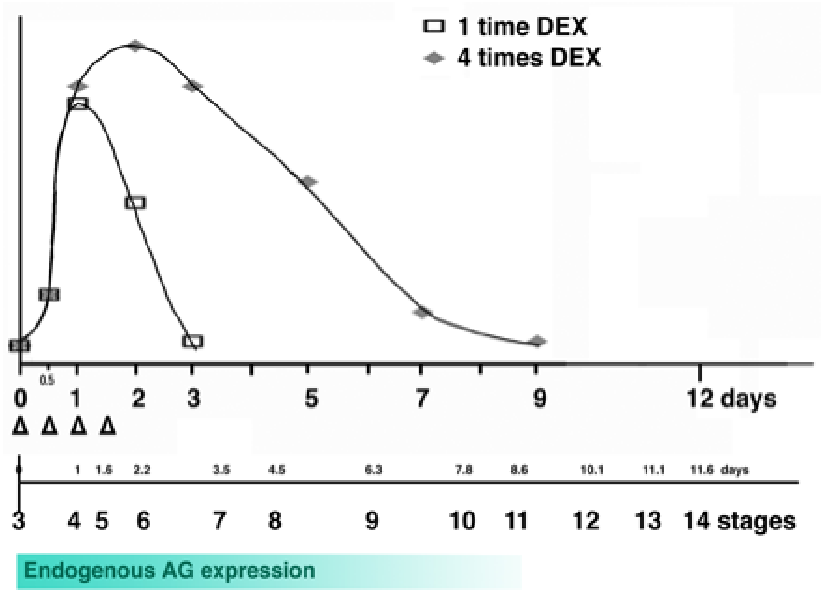


Figure 10: Extended Localization of AG-GR Fusion Protein in Nucleus

DAD1. AG putatively binds to the near-consensus site on the 5' promoter region of DAD1, identified by genomic analysis, resulting in elevated levels of DAD1 expression that ultimately causes anther dehiscence.

On the other hand, MYB26 has two near-consensus AG binding sites, but the timed-induction expression assay gave no proof that the regulatory gene is a direct downstream target of AG. However, MYB26 was shown to be expressed earlier than DAD1 even though it also directly controls anther dehiscence. Therefore, it is probable that MYB26 is involved in an AG-dependent late-stage developmental pathway. Results also provided the evidence that OPR3, COI1, and RBR1 are not transcriptionally regulated due to uniform levels of expression.

5.2 Necessity of Prolonged AG Induction for Flower Development

On a more general level, prolonged AG activity was found to be necessary for late-stage development of normal stamens, although endogenous expression levels did decrease over time, which implies dependence on more complex hierarchies involving indirect targets in late-stage development. AG was also phenotypically observed through the inducible system to be required for anther dehiscence. Therefore, the molecular and phenotypic evidence illustrated the direct control of late-stage development by AG through a regulatory hierarchy. In turn, the results also support the second hypothesis that AG directly and continually regulates the expression of many genes with different functions that act throughout reproductive-organ development.

5.3 Future Work

Although the research expanded on the developmental model, much more work needs to be devoted to revealing the relationship and functions of these regulatory genes. Results from the GUS staining of the plasmid constructs, which should come in within a couple months, will more clearly indicate whether these downstream genes are direct or indirect targets of AG *in vivo*, in addition to determining their spatial expression. The downstream genes that were classified as AG-dependent by current results could still turn out to be direct targets since samples will be analyzed individually. Chromatin immunoprecipitation will also be conducted to observe the physical binding of AG to the identified binding sites. Finally, genome-wide microarray analysis can be conducted to identify more downstream genes of AG that regulate late-stage development since the current research has illustrated how a plant homeotic protein functions throughout development.

5.4 Acknowledgments

I would like to thank my mentors, Dr. Toshiro Ito, Senior Research Scientist, and Professor Meyerowitz, Division of Biology Chair, for their continued support, guidance, and inspiration. I would also like to thank my tutor, Jacqueline Choi, and fellow students, staff, and lab members for their invaluable help with the research paper and presentation. I greatly appreciate the opportunity provided by the Center for Excellence in Education to conduct research through the 2004 Research Science Institute.

References

[1] H. Yu, T. Ito, Y. Zhao, J. Peng, P. Kumar, and E.M. Meyerowitz. Floral homeotic genes are targets of gibberellin signaling in flower development. Proceedings of the National Academy of Science 101 (2004), no. 20, 7827–7832.

[2] S. Harris. *Arabidopsis*—map makers of the plant kingdom. Available at `http://www.nsf.gov/od/lpa/news/publicat/nsf0050/arabidopsis/arabidopsis.htm` (2004/07/30).

[3] R.J. Scott, M. Spielman, and H.G. Dickinson. Stamen structure and function. The Plant Cell 16 (2004), S46–S60.

[4] E. Meyerowitz. The genetics of flower development. Scientific American (November 1994), 56–65.

[5] T. Jack. Molecular and genetic mechanisms of floral control. The Plant Cell 16 (2004), S1–S17.

[6] S. Ishiguro, A. Kawai-Oda, J. Ueda, I. Nishida, and K. Okada. The *DEFECTIVE IN ANTHER DEHISCENCE1* gene encodes a novel phospholipase A1 catalyzing the initial step of jasmonic acid biosynthesis, which synchronizes pollen maturation, anther dehiscence, and flower opening in *Arabidopsis*. The Plant Cell 13 (2001), 2191–2209.

[7] A. Stintzi and J. Browse. The *Arabidopsis* male-sterile mutant, opr3, lacks the 12-oxophytodienoic acid reductase required for jasmonate synthesis. Proceedings of the National Academy of Science 97 (2000), no. 19, 10625–10630.

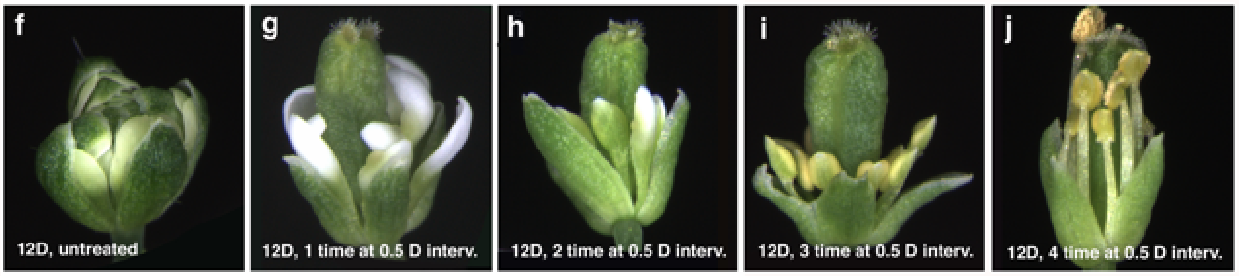


Figure 11: Timed-Induction Floral Phenotypes

Figure 12: AG-Induced (4 × DEX) Inflorescences in Different Stages

AGAMOUS

SPOROCYTELESS/NOZZLE Regulatory Genes Regulatory Genes DAD1 MYB26

Microsporogenesis Filament Elongation Anther Morphogenesis COI1, OPR3

Anther Dehiscence

Figure 13: The "continued" hypothesis of AG function in late-stage development

[8] B.J.F. Feys, C.E. Benedetti, C.N. Penfold, and J.G. Turner. *Arabidopsis* mutants selected for resistance to the phytotoxin coronatine are male sterile, insensitive to methyl jasmonate, and resistant to a bacterial pathogen. The Plant Cell 6 (1994), 751–759.

[9] D. Xie, B.F. Feys, S. James, M. Nieto-Rostro, and J.G. Turner. *COI1*: An *Arabidopsis* gene required for jasmonate-regulated defense and fertility. Science 280 (1998), 1091-1094

[10] S. Steiner-Lange, U.S. Unte, L. Eckstein, C. Yang, Z.A. Wilson, E. Schmelzer, K. Dekker, and H. Saedler. Disruption of *Arabidopsis thaliana MYB26* results in male sterility due to non-dehiscent anthers. The Plant Journal 34 (2003), 519–528.

[11] J. Riechmann. *The* Arabidopsis *book*. The American Society of Plant Biologists, Rockville, MD (2002).

[12] T. Ito, F. Wellmer, H. Yu, P. Das, N. Ito, M. Alves-Ferreira, J.L. Riechmann, and E.M. Meyerowitz. The homeotic protein AGAMOUS controls microsporogenesis by regulation

of SPOROCYTELESS. Nature 430 (2004), 356–360.

[13] T. Ito, H. Sakai, and E.M. Meyerowitz. Whorl-specific expression of the SUPERMAN gene of *Arabidopsis* is mediated by *cis* elements in the transcribed region. Current Biology 13 (2003), no. 17, 1524–1530.

[14] H. Yu, T. Ito, F. Wellmer, and E.M. Meyerowitz. Repression of AGAMOUS-LIKE 24 is a crucial step in promoting flower development. Nature Genetics 36 (2004), no. 2, 157–161.

Design and Characterization of a Littrow Configuration External Cavity Diode Laser

Wenxian Hong

under the direction of
Professor Oskar J. Painter
California Institute of Technology

Abstract

An enhanced tunable Littrow-configuration external-cavity diode laser (ECDL) is constructed from inexpensive commercial components and then characterized. The introduction of an additional mirror parallel to the diffraction grating allows the laser to be tuned without changing the direction of the output beam. The 785 nm ECDL has a typical power output of 4 mW at 56.0 mA of current input with an effective tuning range of 13 nm and exhibits excellent output-beam directional stability. An upper limit on its linewidth is placed at 63.0 MHz.

1 Introduction

External-cavity diode lasers (ECDL) have widespread applications in optical and atomic physics. Such lasers make use of inexpensive yet efficient semiconductor diode lasers which are pumped by electrical energy [4]. The diode is coupled to an external cavity that incorporates a diffraction grating as a wavelength-selective element, which then provides frequency-selective optical feedback to the diode laser via its antireflection-coated output facet. This concept of frequency selective feedback allows the laser to achieve narrow linewidth and remarkable tunability. Typically, such feedback is obtained through diffraction gratings in either the Littrow or Littman-Metcalf configurations.

In the more popular Littrow configuration, shown in Figure 1, the grating is aligned such that the first-order diffraction from the grating is coupled directly back into the laser while the zeroth-order diffraction is reflected as the output beam. The lasing wavelength is dependent on the angle of the incident laser beam with respect to the grating, otherwise known as the Littrow angle θ. Thus, by changing the Littrow angle, precise wavelength tuning can be achieved. The Littrow configuration offers the advantage of high efficiency and power but also faces problems of mode-hopping and beam angular displacement as the grating angle is adjusted.

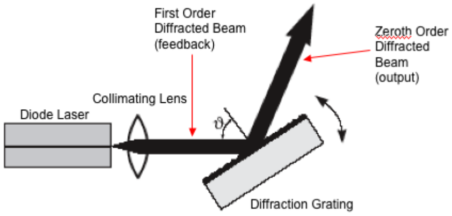


Figure 1: Schematic diagram of external cavity in Littrow configuration. Adapted from [10]

On the other hand, the Littman-Metcalf configuration [8, 9], as depicted in Figure 2, utilizes a diffraction grating at grazing incidence. The first-order diffraction beam goes to an additional mirror, which then reflects the beam back to the grating and into the diode laser as optical feedback. Tuning is achieved by varying the mirror angle instead of the grating angle, allowing the grating, and thus the zeroth-order output beam, to remain fixed as wavelength is changed. Relative to the Littrow design, the Littman-Metcalf configuration overcomes the problems of mode-hopping and beam angular displacement but does not share as high an output efficiency. Furthermore, the additional components involved and its design complexity prevent it from surpassing the widespread prevalence of the relatively simpler yet adequately effective Littrow arrangement.

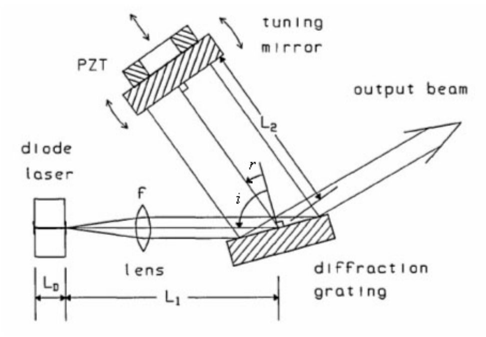


Figure 2: Schematic diagram of an external-cavity diode laser in the Littman-Metcalf configuration [7]

Here, an external-cavity diode laser is constructed based on the Littrow configuration, in close reference to the design of Hawthorn, Weber, and Scholten [6]. Minor modifications have been made to circumvent the problem of beam angular displacement and to produce a fixed-direction output beam. With a fixed output beam, the flexibility of the ECDL is enhanced as the output beam is now able to undergo fiber coupling for delivery to distant applications. The power output and linewidth of the modified ECDL are then determined using instruments such as an optical spectrum analyzer (OSA), a scanning Fabry-Perot interferometer (SFP), and a Michelson interferometer. This modified design is not only low-cost and simple to build, but also produces single-mode emissions of varying wavelengths sufficient for most applications, particularly pertaining to high-resolution, broadband spectroscopy.

2 Design and Construction

The design of the external-cavity diode laser presented here takes advantage of a simple modification of standard Littrow-configured ECDLs to produce a fixed-direction output beam without significant output-power loss. The ECDL, which was named Promise (Figure 3), consists of a 70 mW, 785 nm laser diode (Sanyo DL7140-201S) and aspheric collimating lens (Thorlabs C230TM-B, $f = 4.5$ mm and 0.55 NA) mounted in a collimation tube (Thorlabs LT230P-B). The threaded collimation tube allows for accurate focusing of the collimating lens and also aligns the lens axis with the laser diode axis. The diode-tube assembly is then fixed to the back plate of a modified optical mount (Newport U100-P2K). The modifications on the optical mount include the tapping of three holes into the front plate and the removal of a square section of the plate. The diffraction grating used (Thorlabs GR13-1208) is gold-coated and blazed at 28° with 1200 lines/mm on a 12.7 mm × 12.7 mm × 6 mm substrate. A grating mount is constructed to hold the diffraction grating and is subsequently attached to the front plate of the modified U100-P2K. The grating is vertically and horizontally adjusted by the mirror mount actuators for coarse wavelength tuning.

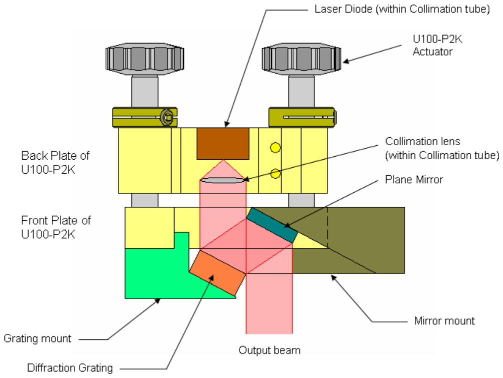


Figure 3: Schematic diagram of the enhanced Littrow-configuration external-cavity diode laser (Promise), as viewed from above

The main modification made to ensure a fixed-direction output beam is the addition of a single plane mirror (Newport 05SD520ER.2) parallel to the tuning diffraction grating and located rigidly with respect to it. This is done by attaching the mirror to a mirror mount that is then secured to the front plate of the U100-P2K. The laser beam reflects from the grating and then from the mirror, thus replacing an otherwise large angular beam displacement with a negligible translational beam displacement. When the grating is rotated by $\Delta\theta$, the beam reflected by the grating is rotated by twice that angle, or $2\Delta\theta$, by the law of specular reflection. Since the mirror rotates by the same angle as the grating, when the reflected beam from the grating strikes the mirror, it is again reflected back by $2\Delta\theta$, such that the output beam remains relatively fixed. This concept of adding a mirror to prevent beam displacement has also been employed in the design of Hawthorn, Weber and Scholten [6] and is an improvement on the previous design by Arnold, Wilson, and Boshier [2]. However, due to the differences in blaze angle of gratings used, minor design changes were made to the positions of the mirror and grating so that they were parallel and did not obstruct the output beam.

The compact nature of the Promise ECDL also provides for simple and efficient temperature regulation of the entire system using a Peltier thermoelectric cooler (Melcor CP1.4-71-045L). A 10 kΩ thermistor (Thorlabs TH10K) attached to the collimation tube assembly provides feedback to the thermoelectric cooler via a temperature controller, thus ensuring reasonable thermal stability. Thermal paste is applied at each interface to ensure optimal thermal conductivity and efficient cooling of the setup. In addition to the thermistor used for temperature control, another temperature sensor (LM35) on the collimation tube gives independent readouts. Other improvements made to the ECDL system include the use of a massive aluminum metal base/heat sink to provide inertial and thermal damping. The ECDL setup is isolated from the optical table by placing a layer of Sorbothane beneath the baseplate as well as by enclosing the laser with an aluminum cover, which is also isolated from the laser by Sorbothane. This dampens acoustic noise and shields the laser from air currents, thus providing for great stability and precision of the instruments.

The design of Promise described above calls for the modifications of several components and the construction of other new parts. During the time spent waiting for the machining of components, a second ECDL, named Prowl, was constructed using existing components. The concept of the latter laser remains the same as the former, making use of the modified Littrow configuration, which introduces the plane mirror that moves simultaneously with the rotation of the diffraction grating. The only visible difference is that instead of containing the entire ECDL system within a single U100-P2K mount, the Prowl design makes use of the back plate of one U100-P2K mount to hold the laser diode and another unmodified U100-P2K mount to provide for manual translation of the grating and mirror. All essential aspects of the laser, such as the thermoelectric cooler, have been included, though less important components such as the independent temperature sensor have been omitted. These minor modifications will only serve to lengthen the cavity length slightly, reducing the mode spacing of the external cavity, but should not significantly hinder the performance of the laser. Thus, the ECDL design of Prowl is likely to demonstrate the strengths of the modified Littrow configuration to as great an extent as the initially intended design of Promise.

3 Methodology

Upon construction, the lasers were characterized based on certain performance indicators, namely output power, tuning range, output-beam directional stability, and linewidth. A battery of tests was run using various devices to analyze the laser performance, the methodologies of which will be presented in this section. However, due to time constraints, the experiments outlined mainly pertain to the performance characteristics of Prowl. Nevertheless, similar experiments can be conducted with the Promise design to evaluate its effectiveness when time permits.

3.1 Alignment of Feedback

The key to the Littrow configuration is the back-coupling of the first-order diffraction beam from the grating into the laser diode. Without this feedback, the Littrow laser cannot achieve single-mode emission and will lase at a wavelength set by the gain peak of the semiconductor active region. Thus, it is critical to first align the feedback beam before carrying out further experiments. Preliminary alignment is done visually by rotating the actuators and merging the secondary feedback and output beams into a single bright spot on a screen at the output of the laser. Good alignment is confirmed by inspecting the laser spectrum using an optical spectrum analyzer or scanning Fabry-Perot interferometer, and modifications are made until single-mode emission is obtained. Aligning the feedback beam is also considerably easier when the laser is close to its lasing threshold, since the laser is most sensitive to feedback at this point. When the feedback beam is well aligned, interference fringes are expected in the output beam and the laser remains single-mode with temperature and current changes.

3.2 Optical Spectrum Analyzer

An optical spectrum analyzer is an instrument that displays the wavelength-dependent intensity of input light. This is achieved by means of a rotating diffraction grating that deflects light of different wavelengths at different angles. A fixed detector measures the intensity of incident light at a given angle, thus determining the wavelength-dependent intensity as the diffraction grating is rotated. Using this instrument, we are able to determine the effective tuning range of the Prowl ECDL as well as its output power.

The output beam of the Prowl ECDL was directed to a fiber coupler leading to the OSA located 70 cm from the output mirror of the laser. The ECDL actuators were rotated so that the diffraction grating was tilted about an axis parallel to its grooves, thereby giving rise to coarse wavelength tuning. The laser wavelength was continuously tuned until a significant mode-hop back to the central gain peak was detected, signaling that the wavelength-dependent gain had overcome the wavelength-selective influence of the feedback. This demarcated the upper and lower limits of the ECDL tuning range. Five digital screen shots were taken across the entire tuning range, spaced at approximately equal wavelength intervals. The linewidth of the laser can also be approximated from a single normalized digital scan. Throughout all these experiments, the temperature was regulated by the thermoelectric cooler at a thermistor resistance of 11.0 kΩ, corresponding to a relatively constant temperature of 23.9°C, so as to minimize the adverse effects of thermal drift.

3.3 Scanning Fabry-Perot Interferometer

The scanning Fabry-Perot interferometer provides a greater resolution for the linewidth measurement as compared to the OSA. The SFP comprises a regular Fabry-Perot etalon with a fixed mirror arm and a piezoelectric actuator attached on the other partially transmitting mirror arm. As voltage is applied across the piezostack, the Fabry-Perot cavity length constantly changes by a small displacement and thus varies the standing wavelengths within the cavity. A detector measures the intensity of light that eventually leaves the cavity via the partially transmitting mirror, while an oscilloscope displays the intensity of light against the voltage applied to the piezostack, which can later be related to frequency.

The SFP used in the characterization of Prowl is the Tropel Model 240 Spectrum Analyzer, with a finesse of 200 at 700 nm and a free spectral range (FSR) of 1.5 GHz. The output beam from the Prowl ECDL is directed through the SFP to obtain a repeating series of Lorentzian lineshape functions. After the normalization of waveforms, the linewidth of a single wave is then determined. The SFP is also particularly useful in ascertaining the quality alignment of the grating feedback, as multi-mode emission lines can be clearly resolved.

3.4 Michelson Interferometer

In a Michelson interferometer, a beam splitter divides the incident waves from a light source into two samples, which are then directed along separate paths. Each beam of light travels a different path length along each arm of the interferometer, acquiring a relative phase difference in the process. When the beams are reflected back to the beam splitter, they recombine and interfere to produce a interference pattern. By examining the interference in the recombined beam, information on the source spectrum can be obtained. Here the Michelson interferometer is employed to observe the visibility of fringes produced at the output beam. As the relative path length is lengthened by moving one of the arms away from the beam splitter, the beam in the arm gradually loses coherence, causing the fringes to oscillate between constructive and destructive interference and leading to an overall decrease in fringe visibility. A photodetector placed at the recombined beam measures the intensity of the fringes and translates the data to an oscillating voltage output. The envelope of the oscillating output is then displayed on an oscilloscope.

This decay in fringe visibility occurs as a result of broadening effects, in particular Lorentz broadening. Lorentz broadening (or lifetime broadening) results from species in excited states having finite lifetimes. From Heisenberg's uncertainty principle, the product of uncertainties in energy levels and lifetimes is on the order of $\hbar$. Hence, the finite lifetime of an excited species leads to a spectrum of energy emitted when the species returns to ground state, giving rise to broadening of the laser waveform. The fringe visibility function, $V(\tau)$, for a Lorentz broadened wave is described by the exponential curve

$$V(\tau) = \exp -\pi\tau\Delta\nu/2, \tag{1}$$

where τ is the time difference between the two paths and $\Delta\nu$ is the linewidth of the laser mode [3].

The Michelson-interferometer measurements were carried out with a current input of 40.1 mA and a thermistor resistance of 11.0 kΩ. One arm of the interferometer was lengthened by 45 cm using a Unislide carriage, and the fringe visibility curve was plotted by the oscilloscope. Three repetitions were carried out with the same interferometer alignment configuration and with the Unislide moving at the same velocity. By fitting the empirical data obtained from the oscilloscope on a similar exponential, the linewidth of the laser can then be calculated.

4 Performance

4.1 Tuning range of ECDL

The digital screenshots taken from the OSA were compiled using MATLAB to describe the tuning range of the Prowl ECDL. At 56.0 mA, the wavelength of laser output can be tuned over a 11 nm range from 775 nm to 786 nm by rotating the grating alone. At lower current inputs of 37.8 mA, wavelength tunability can be further extended to a wider range of 13 nm between 774 nm and 787 nm, as shown in Figure 4. A qualitative explanation of this phenomenon stems from the knowledge that the wavelength at which spontaneous emission occurs is dependent on the gain spectrum, while the lasing wavelength is dependent on both the gain spectrum and reflectivity of the grating. Thus, at higher currents, the broad gain spectrum tends to overwhelm the influence of the narrow band grating reflectivity, making it more difficult for the ECDL to lase based on frequency-selective feedback. However, when the current input nears the lasing threshold limit at 33.0 mA (Figure 5), the tuning range of the ECDL once again shrinks slightly to 12 nm (774–786 nm). This is because the amplitude of the gain spectrum decreases with lower currents. Thus, the range of wavelengths at which gain exceeds loss (i.e., possible lasing wavelengths) is reduced, limiting the tuning range of the ECDL.

It is also worthwhile to note that the gain spectrum of the ECDL, although approximated by a Gaussian in Figures 4 and 5, is actually slightly asymmetrical. This is because the semiconductor material in the laser diode absorbs at shorter wavelengths compared to the wavelengths at which it emits. Since gain is proportional to the difference between emission and absorption of photons, higher gain is obtained at longer wavelengths, resulting in a steeper gradient on the gain spectrum at shorter wavelengths and an asymmetrical gain spectrum shape.

These wavelength measurements do not take into account the effect of temperature variations, which should allow the laser to tune over an even wider range. Nevertheless, the wavelength-tunability values obtained for the Prowl ECDL are reasonably comparable to commercial standards, demonstrating the feasibility of the design.

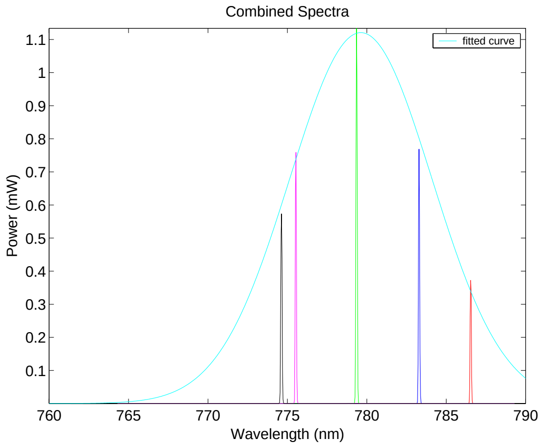


Figure 4: OSA plot of intensity against wavelength at 37.8 mA, showing the tuning range of the ECDL. The fitted curve represents the projected gain spectrum of the laser.

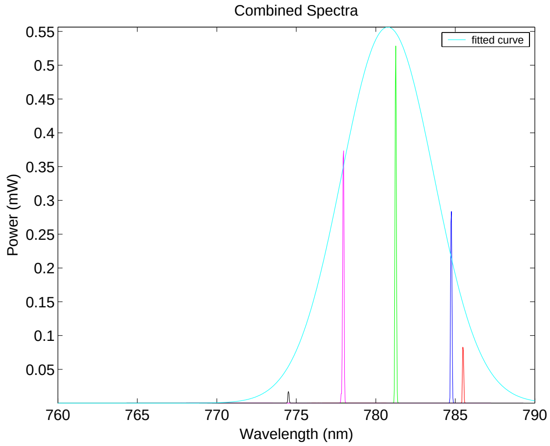


Figure 5: OSA plot of intensity against wavelength at 33.0 mA, showing the tuning range of the ECDL. The fitted curve represents the projected gain spectrum of the laser.

4.2 Output power of ECDL

With the laser wavelength tuned to the gain peak, the output power of the Prowl ECDL was measured to be approximately 1.1 mW at 37.8 mA of current input and about 4.0 mW with 56.0 mA of current. It may seem surprising that the power output of the Prowl ECDL is a mere fraction of the typical power output of the 70 mW laser diode. However, this is expected since the diffraction grating used has a high efficiency of 60%–80% at blaze wavelength. Consequently, most of the diode power output will be reflected back into the diode by the grating instead of being directed into the zeroth-order beam, resulting in the observed low power output. However, the small power coupled out of the ECDL is sufficient to accurately characterize the laser and should not pose too great a concern.

4.3 Output-beam directional stability of ECDL

The output-beam directional stability of the Prowl ECDL was demonstrated by monitoring the tuning of wavelength as the grating is rotated using an OSA located 70 cm from the laser. Despite the high sensitivity of the OSA fiber coupler to beam misalignment, the full tuning range of 13 nm was scanned successfully by adjusting the grating angle, without any need for realignment. This wavelength variation will normally displace the output direction of the laser beam by about 0.45° and cause a lateral translation of 5.5 mm at the fiber coupler. However, with the additional mirror, the lateral displacement of the output beam for a small change $\Delta\theta$ in grating angle is now given by $\Delta x \sim 2L\Delta\theta$, where L is the distance the beam travels between the grating and mirror.

Given that L (~ 3 cm) is far smaller than 70 cm, the geometric ingenuity of the Prowl design has not only enabled the angle of output beam to remain unchanged, but also made the lateral displacement negligible at approximately 0.47 mm. Such minute shifts are insignificant for most applications, indicating the success of the Prowl design in maintaining a fixed-direction output beam.

4.4 Linewidth of ECDL

An attempt to measure the linewidth of the Prowl ECDL was first made using the OSA. A single Lorentzian lineshape waveform obtained at the laser emission wavelength at 60.0 mA was plotted and normalized. By examining the full width half maximum (FWHM) of the waveform, the linewidth of Prowl was measured to be approximately 0.08 nm, or 39 GHz, at 60.0 mA (Figure 6), which corresponds to the resolution bandwidth of the OSA. Intuitively, this value is likely to be too large as commercial lasers of the same standard generally have linewidths in the order of kHz or MHz. Hence, it is reasonable to infer that the OSA lacks the necessary resolution bandwidth to resolve the narrow linewidth of the laser.

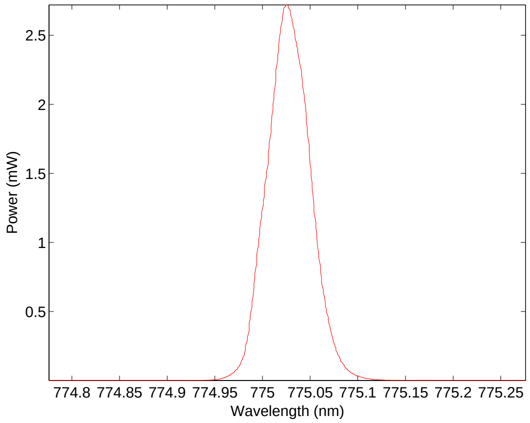


Figure 6: Measurement of linewidth using the optical spectrum analyzer. Note that the linewidth is limited by the resolution bandwidth of the OSA (~ 0.08 nm).

Therefore, the SFP was employed to determine the linewidth of Prowl to greater precision. After single-mode emission was obtained, two adjacent peaks were normalized and spaced at nine divisions apart by adjusting the x-axis scale of the oscilloscope (Figure 7a). Since the FSR of the SFP is given as 1.5 GHz, the voltage separation on the x-axis between the peaks can now be related to frequency. The FWHM is then measured on the normalized scale. Figure 7b shows that at 50 mA, the FWHM of the waveform is measured to be approximately 0.7 of a division. Using simple proportion calculations, the FWHM corresponds to about 120 MHz. At 70 mA, the FWHM is exactly 1 division long, or 170 MHz in terms of frequency.

Now, it is imperative to note that the FWHM measurements obtained could refer to either the Prowl ECDL linewidth or scanning Fabry-Perot linewidth. In order to distinguish between the two, the effect of tilting the SFP on the waveform obtained is examined. By tilting the SFP slightly, the reflectivity of the mirrors drops sharply, resulting in a significantly lower finesse. Recalling the relationship $\Delta\nu = \nu_{\mathrm{FSR}}/F$, a lower finesse will give rise to a greater Fabry-Perot linewidth. So, if the Fabry-Perot linewidth were much narrower than the

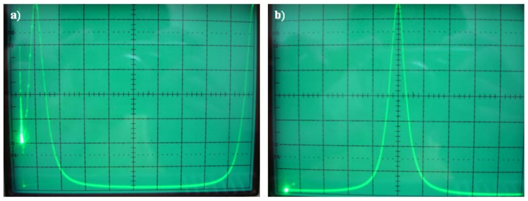


Figure 7: a) Normalization of adjacent peaks on SFP oscilloscope. b) Measurement of linewidth on normalized scale using the SFP.

laser linewidth, any slight increase in the Fabry-Perot linewidth should still not alter FWHM measurements significantly. However, it is indeed observed that by changing the angle of incident light on the SFP, the FWHM measurements do not remain constant. Therefore, one can infer that the Fabry-Perot linewidth must be bigger than that of the laser. As such, even though the SFP measurements have narrowed down the laser linewidth by a factor of 300, this is still insufficient to totally resolve the narrow laser linewidth.

This result is not surprising, as the Tropel Model 240 Spectrum Analyzer is designed for wavelengths ranging from 600 nm to 700 nm, which is not within the wavelength range of the ECDL ($\sim$ 780 nm). The dielectric mirrors of the SFP have a narrow bandwidth of wavelengths for which they have high reflectivity, and beyond this bandwidth, the finesse of the SFP drops drastically (a value of 100 MHz for the Fabry-Perot linewidth would actually imply that the finesse has fallen from 200 to 15). In order to provide greater resolution for the laser linewidth, a SFP with a higher finesse at the required wavelength should be used. However, the measurements of the SFP have served a useful purpose in placing an upper boundary on the linewidth of the laser, meaning that the linewidth cannot exceed frequencies on the order of 100 MHz.

The linewidth measurements obtained from the Michelson interferometer experiment were slightly more satisfactory. Assuming that the decay curve obtained from the oscilloscope follows the exponential in Equation (1), we are able to describe the exponential curve using the initial voltage V_0 and the final voltage V_f. Equation (1) then changes to

$$V_0 \exp -\pi\tau\Delta\nu/2 = V_f. \tag{2}$$

Solving for $\Delta\nu$ yields

$$\Delta\nu = -\frac{2}{\pi\tau}\ln\frac{V_f}{V_0}. \tag{3}$$

Substituting the relevant values of τ, V_0 and V_f into Equation (3), allows the linewidth of the laser to be easily calculated. Table 1 shows the results of the Michelson-interferometer experiments conducted and the linewidth values obtained.

τ / ns	V_0 / V	V_f / V	$\Delta\nu$ / MHz
3.0	11.6	8.1	76.2
3.0	11.1	7.6	80.4
3.0	10.9	7.7	73.8

Table 1: Summary of Michelson interferometer experiment.

By averaging the three linewidth values obtained, the average linewidth of the laser is calculated to be 76.8 MHz. This is a conservative estimate, but nevertheless further narrows the linewidth range of the laser. However, undulations observed in the empirical decay curve have rendered the linewidth values less accurate than desired. This problem can be overcome by only considering part of the decay graph, where the progression seems relatively exponential. Given that the average velocity of the Unislide carriage was 0.20 $\mathrm{cm \cdot s^{-1}}$, and that the timeframe considered was 70 s, a new value of τ was calculated. Table 2 shows the results of the linewidth calculations under these new conditions.

τ / ns	V_0 / V	V_f / V	$\Delta\nu$ / MHz
0.93	8.8	8.1	56.7
0.93	8.4	7.6	68.5
0.93	8.4	7.7	63.9

Table 2: Linewidth calculations after additional considerations on the exponential nature of the data curve. Note that the distance travelled by the Unislide carriage is no longer 45 cm.

This time, by averaging the three linewidth values, a narrower average linewidth of 63.0 MHz is obtained. Considering the mechanical noise produced by the Unislide carriage while in motion, which may disturb the laser, as well as the limited ability of such Michelson interferometer experiments to measure linewidth, a linewidth value below 100 MHz demonstrates the success of this particular experiment. In conclusion, after using the three different instruments, OSA, SFP, and Michelson interferometer, the Michelson interferometer proved to be most precise, and the final upper bound on the linewidth was set at 63.0MHz.

5 Discussion

Prowl has clearly proven itself to be feasible for general use as a tunable laser, from its wide tuning range, mechanical stability and narrow linewidth. One of the biggest advantages of the design is that it primarily makes use of inexpensive, readily available optical components. The total cost of construction was under US$1000, comparatively cheaper than most commercial Littrow-configured ECDLs. Moreover, the ready accessibility of the different parts makes the laser especially simple to construct. The additional mirror introduced has also worked superbly to allow for coarse tuning without alteration of alignment, greatly enhancing the versatility of the conventional Littrow ECDL. Both the Promise and Prowl lasers will be used subsequently in experiments pertaining to atomic spectroscopy and laser cooling and have helped pave the way for other novel designs of compact ECDLs.

A significant problem faced in the construction of the Prowl laser, however, was the ease with which the laser diode could be blown. This occurred despite the use of a current controller that protects the laser diode with a shunting circuit by dissipating current transients when the power cable is detached. The blowing of the diode can be attributed to the loose fit between the diode ends and diode socket. When the diode socket was accidentally disconnected, there was no protective circuitry to prevent strong current transients from entering the diode and ruining its intrinsic material properties. Thus, future design considerations should incorporate a firm coupling to give a tighter fit between the diode and the socket.

Several other improvements can be made to the existing Prowl or Promise design to further extend its applications. These include the installation of a piezostack on the front plate of the U100-P2K and a piezodisk on the diffraction grating, so as to enable fine continuous wavelength tuning across the cavity mode spacing. Secondly, to increase the existing power output, a diffraction grating with lower efficiency at blaze wavelength could be used so that more power is coupled into the output beam instead of the feedback beam. Alternatively, the grating could be deliberately misplaced away from its blaze angle. Lastly, smoother tuning and wider wavelength variations can be achieved by applying an antireflection (AR) coating to reduce the reflectivity of the laser diode output facet. External-cavity lasers using optimally AR coated diodes have achieved continuous tuning ranges of several THz [1]. Thus, an appropriate AR coating will appreciably enhance the tunability of the ECDL.

Future studies could begin with the characterization of the Promise design, since time constraints have prevented the construction and testing of this laser. A comparison could be made between the designs of Promise and Prowl to evaluate the strengths and weaknesses of both lasers. Furthermore, three methods of linewidth measurement have been presented, although these are not as accurate as desired. Higher linewidth resolution could be provided by heterodyning and photomixing methods, where the beat frequency of two ECDLs operating at nearly the same frequency is measured. As an extension, both lasers can be locked and stabilized by the Zeeman effect and the radiofrequency (rf) beat between them observed with an avalanche photodiode [12].

6 Conclusion

The design of a compact tunable Littrow-configured external-cavity diode laser was described. Empirical data from performance tests indicated that the Prowl ECDL has an effective tuning range of 13 nm, a typical output power of 4 mW and a linewidth of at most 63.0 MHz. The laser also exhibited exceptional output-beam directional stability due to the introduction of a plane mirror parallel to the grating. These superior characteristics of Prowl testify to the advantages of the enhanced ECDL design. The simple modifications presented here can be applied to existing Littrow-configuration ECDLs, and the design holds exciting prospects for future applications in atomic and optical physics.

7 Acknowledgments

I would like to extend my sincerest thanks to my mentor, Professor Oskar J. Painter of the California Institute of Technology, and his graduate students Orion Crisafulli and Raviv Perahia for the invaluable guidance and assistance they provided in the course of the project. Also, I would like to thank my tutor, Dr. John Rickert, and the staff of the Research Science Institute at Caltech for their input and suggestions as well as help rendered in times of need. Finally, I would like to express my heartfelt gratitude to the Center for Excellence in Education and the Ministry of Education, Singapore, for giving me this unique opportunity to enrich myself and expand my horizons.

References

[1] A. Andalkar, S.K. Lamoreaux and R.B. Warrington. Improved external cavity design for cesium D1 (894 nm) diode laser. Review of Scientific Instruments 71 (2000), no. 11, 4029–4031.

[2] A.S. Arnold, J.S. Wilson and M.G. Boshier. A simple extended-cavity diode laser. Review of Scientific Instruments 69 (1998), no. 3, 1236–1239.

[3] M. Born and E. Wolf. *Principles of Optics.* 7th ed. Cambridge University Press, Cambridge, United Kingdom (2002).

[4] L.A. Coldren and S.W. Corzine. *Diode Lasers and Photonic Integrated Circuits.* John Wiley & Sons, Inc., New York, NY (1995).

[5] T.M. Hard. Laser wavelength selection and output coupling by a grating. Applied Optics 9 (1970), no. 8, 1825–1830.

[6] C.J. Hawthorn, K.P. Weber and R.E. Scholten. Littrow configuration tunable external cavity diode laser with fixed direction output beam. Review of Scientific Instruments 72 (2001), no. 12, 4477–4479.

[7] K.C. Harvey and C.J. Myatt. External-cavity diode laser using a grazing-incidence diffraction grating. Optics Letters 16 (1991), no. 12, 910–912.

[8] M.G. Littman and H.J. Metcalf. Spectrally narrow pulsed dye laser without beam expander. Applied Optics 17 (1978), no. 14, 2224–2227.

[9] K. Liu and M.G. Littman. Novel geometry for single-mode scanning of tunable lasers. Optics Letters 6 (1981), no. 3, 117–118.

[10] O.I. Permyakova, A.V. Yakovlev, and P.L. Chapovsky. Simple external cavity diode laser. To be published (2003).

[11] B.E.A. Saleh and M.C. Teich. *Fundamentals of Photonics.* John Wiley & Sons, Inc., Canada (1991).

[12] C.E. Wieman and L. Hollberg. Using diode lasers for atomic physics. Review of Scientific Instruments 62 (1991), no. 1, 1–20.

The Effect of Chemical Additives on Ice-Spike Formation

Samuel Lederer

under the direction of
Dr. Kenneth Libbrecht
California Institute of Technology

Abstract

Solutions of sodium bicarbonate, calcium chloride, ethanol, acetic acid, and methanol, in concentrations ranging from 10^{-7} to 0.1 molar, were frozen under controlled conditions to measure the dependence of ice-spike formation on chemical impurities in water. The probability of ice-spike formation was plotted against molarity of solution to determine general trends for different solutes. The varying effects of different solutes at the same concentrations suggested that freezing-point depression was unlikely to cause inhibition of ice-spike formation in solutions, and a model based on dendritic sheet formation was proposed.

1 Introduction

Ice spikes, or spicules, are thin, tapered projections of ice occasionally formed during the freezing of standing water, sleet pellets [3], and icicles [11]. They are also found when distilled water is frozen in ordinary household ice-cube trays (Figure 1). This phenomenon has been observed and recorded for nearly a century [5], but its causes remain poorly understood.

A qualitative model of the phenomenon was put forth independently by Bally and Dorsey in the early 20th century and remains the most widely accepted description of this effect [2, 3, 5, 14]. For the case of water confined to a container, the model outlines the following sequence (Figure 2). Surface water nucleates around irregularities where it meets the container wall and freezes inward, eventually leaving a small hole in the surface. Meanwhile, water adjacent to the walls of the container begins to freeze, and the temperature of water in the center of the cube falls

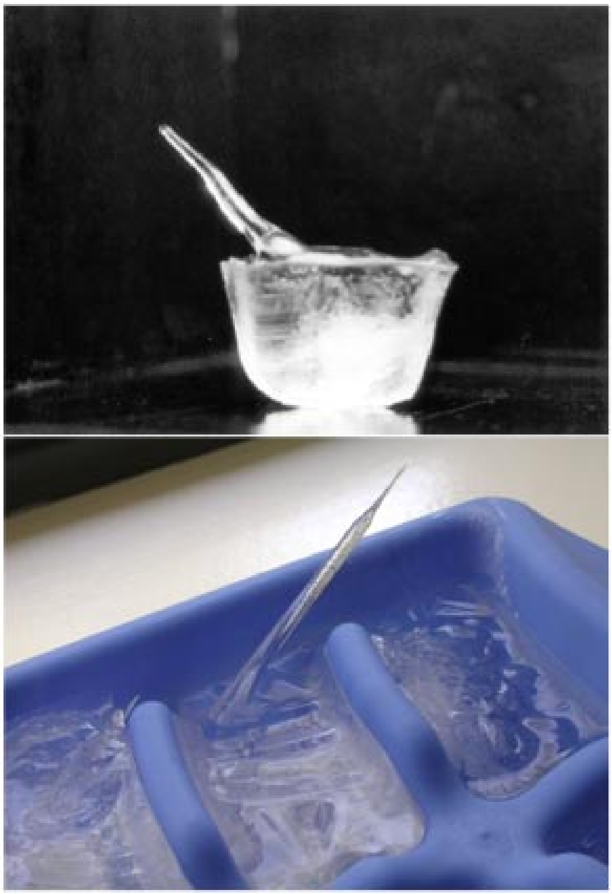


Figure 1: Examples of typical ice spikes from water frozen in standard household trays. The spike protrudes from the surface at an angle and tapers from base to tip. Sources: Morris, Libbrecht [15, 13].

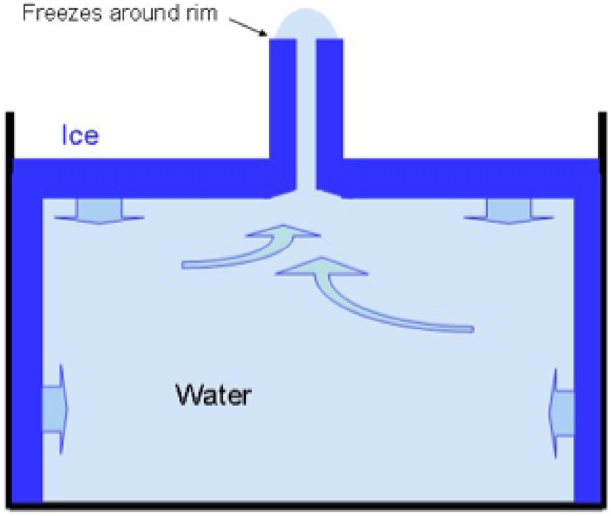


Figure 2: A depiction of the Bally-Dorsey model. Water freezing on the sides of the cube pushes liquid water up through the hole in the surface, where it forms a drop whose freezing edges create the ice spike. Source: Libbrecht, Lui [14].

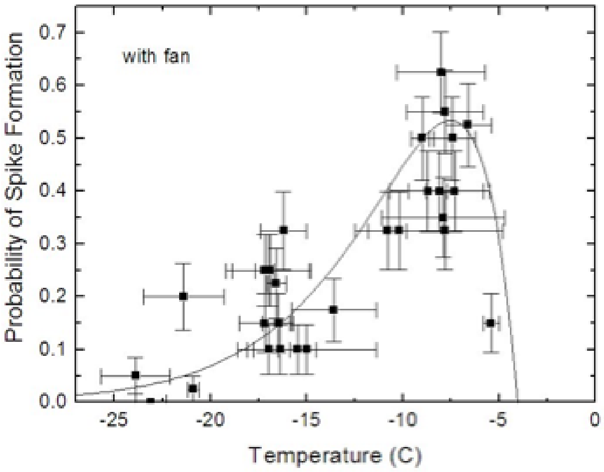


Figure 3: The probability of ice-spike formation versus temperature for pure distilled water in ordinary plastic ice-cube trays. A fan inside the freezing chamber promoted evaporative cooling and facilitated the growth of ice spikes. Probability of ice-spike formation is maximized near −7°C. Source: Libbrecht, Lui [14].

below 4°C, causing it to expand. The expansion of interior water and the reduction of volume in the interior of the cube increase the pressure on the water, pushing it upward through the hole. This results in the formation of a drop of water immediately above the hole. The water at the edges of this drop freezes, initiating the tube. As volume within the cube is increasingly taken up by ice, the remaining liquid water is pushed up through this tube and freezes, lengthening the tube. The growth of the tube ceases when the drop at the top of the tube freezes entirely, which is substantially before the rest of the water in the cube is frozen.

Previous research has shown that the formation of ice spikes in pure distilled water is maximized when the water is frozen in ordinary plastic ice-cube trays at a temperature near −7°C in the presence of adequate air circulation (Figure 3) [14]. Other reports find that the probability of ice-spike formation is drastically reduced by the presence of impurities in the water [11, 13, 14], an effect which is not addressed in the Bally-Dorsey model. In particular, although ice spikes form readily in most home freezers when distilled water is used, they rarely form from tap water. It is possible that this is due to the colligative property of freezing-point depression:

$$\Delta T_f = \frac{-1.86\,\mathrm{kg}\cdot{}^\circ\mathrm{C}}{\mathrm{mol}} \cdot mi, \tag{1}$$

where ΔT_f is the change in freezing point; m is the molality of the given solution; and i is the van't Hoff factor, which describes the dissociation of the solute into component ions. For ethanol, methanol, sodium bicarbonate, and calcium chloride, the i factors are 1, 1, 2, and 3, respectively. For acetic acid, the value of i varies as the inverse square root of concentration, with values between 1 and 2 in this experiment. Note that for low concentrations, such as those involved in this experiment, which were less than or equal to 0.1 molar, molality is numerically equivalent to molarity, so these measures will be used interchangeably in calculations.

Eighty percent of tap-water systems in the United States use groundwater [7], and the maximum levels of chemical impurities mandated by the EPA are negligible in comparison to dissolved minerals [8], so a calculation of the freezing-point depression in groundwater may be used to estimate the freezing-point depression in tap water. A common measure of the presence of these mineral impurities in water is total dissolved solids (TDS), expressed in mg/L.

The EPA advises, but does not require, that TDS not exceed 500 mg/L [6]. Estimating TDS as 1000 mg/L, the approximate concentration of impurities in the water may be obtained by consulting Table 1. The total molality of groundwater is approximately 0.002 mol/kg, which would produce a freezing-point depression of only 0.004°C, yet ice spikes are extremely rare in tap water. However, there are also some reports of regular ice-spike formation in tap water. These facts strongly suggest that factors other than freezing-point depression are involved. The purpose of this research is to investigate which chemical impurities facilitate or hinder ice-spike growth and to determine the variation of probability of ice-spike formation as a function of chemical-impurity concentration.

Component	**% TDS**
Silica (SiO_2)	3.0
Calcium	11.2
Sodium	10.0
Magnesium	2.5
Bicarbonate (HCO_3^-)	62.5
Sulfate (SO_4^{2-})	10.0
Chloride (Cl^-)	5.0

Table 1: Relative concentrations of impurities in groundwater, measured by percentage of total dissolved solids. Source: Ozsvath [16].

2 Materials and Methods

The freezer for this experiment was an insulated copper cylinder cooled by circulating chilled methanol through copper tubing attached to its top and sides. The air temperature within the freezer was maintained near −9°C, and a small electric fan was mounted inside to promote evaporative cooling. Household plastic ice-cube trays were cut to fit inside the tank and, when filled, were positioned on a wire rack approximately 25 cm above the bottom of the tank. Solutions frozen consisted of sodium bicarbonate, ethanol, calcium chloride, acetic acid, and methanol dissolved in laboratory deionized water. Only one solute was present in each trial. Concentrations tested ranged from 10^{-7} to 10^{-1} molar. For each of the trials, 5 trays, each containing 10 cubes, were filled with the appropriate solution, placed in the freezer for approximately 6 hours, and then removed and inspected.

For analysis purposes, an ice spike was defined to be a distinct projection from the ice surface of length 0.5 cm or greater. The total number of cubes with ice spikes was counted, and this number was divided by 50 to estimate the probability of ice-spike formation for the given conditions. Further trials were sometimes necessary and are reflected in reduced error margins on the data plots.

3 Results

3.1 Observations

The experimental data are summarized in Figure 4. Pure water exhibited a probability of ice-spike formation of roughly 0.34. The separate addition of both calcium chloride and sodium bicarbonate in concentrations less than 10^{-5} molar resulted in a slightly higher probability of ice-spike formation, although this may have been purely due to experimental error. With increasing concentrations, the two salts showed similar behavior, drastically reducing the probability of ice-spike formation. These results are consistent with the results of a previous study using sodium chloride [14]. Both methanol and ethanol, added separately to water, significantly facilitated the formation of ice spikes, only producing an inhibitory effect at concentrations greater than or equal to 10^{-3} molar. Acetic acid produced a larger inhibitory effect than other solutes at very low concentrations, but behaved in a similar manner as as the two salts at concentrations greater than 10^{-5} molar.

In addition to the number of ice spikes, the appearance of the frozen cubes was also observed. Two effects were found to occur at high concentrations. The first was a smoothing of the surface of the cube. Cubes consisting of pure, or nearly pure, water usually exhibit a few millimeter-scale lumps distributed randomly over the surface—features which were almost entirely absent in cubes of solutions of impurity concentration greater than 10^{-3} molar. Although no quantitative data were recorded, the appearances of ice lumps and ice spikes seem highly correlated. The second effect was a change in the appearance of the insides of the ice cubes. High concentrations of acetic acid, methanol, and ethanol sometimes produced cubes that were nearly transparent. Pure water cubes exhibited a cloudy appearance due to gases in the water which came out of solution during freezing, leaving bubbles.

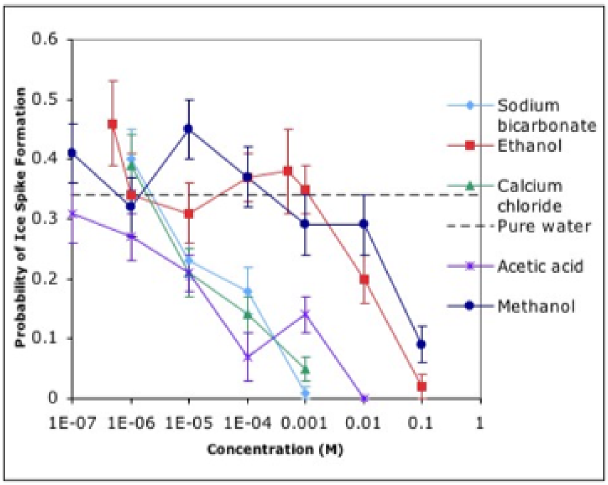


Figure 4: Consolidated experimental data. Concentrations are molar. Data for sodium bicarbonate ($NaHCO_3$), calcium chloride ($CaCl_2$), and acetic acid (CH_3COOH) were similar, producing a large inhibitory effect at relatively low concentrations. Ethanol (CH_3CH_2OH) and methanol (CH_3OH) exhibited no inhibitory properties at low concentrations and required roughly 100 times the concentration of the other solutes to have a similar inhibitory effect.

3.2 Sources of Error

This experiment was designed to determine general trends, so minimizing error was secondary to acquiring substantial amounts of data. Error bars shown in plots are one standard deviation, calculated using Bernoulli statistics. In addition to random error due to vibrations and minute temperature fluctuations, there were several sources of systematic error in this experiment, most of which were considered inconsequential. The first of these was the inaccuracy of the concentrations of the solutions frozen. Volumes were measured in a 300 mL beaker accurate to about 3 mL and in a 2 L beaker accurate to about 10 mL. Masses were measured with an electronic balance accurate to 0.1 g. These instruments may have contributed an error of as much as 10% in the concentrations, which was judged harmless due to the logarithmic distribution of trial concentrations.

There were two main sources of human error. The first was in the identification of ice spikes. The criteria were subjective, and verification of appropriate length was accomplished by eye, without use of a ruler. The second was in the placement of the ice-cube trays on the wire rack, which varied slightly from trial to trial. Both of these sources of error could have been addressed, but this would have required a prohibitive amount of time and reduced the amount of data collected.

4 Discussion

The data prompt many questions about the ice-spike phenomenon. The presence of impurities can clearly inhibit ice-spike formation, and this effect becomes more prominent at higher concentrations. However, a preliminary model for ice-spike inhibition based on freezing-point depression seems unlikely due to the scale of the concentrations involved. For instance, a 0.001 molar solution of $CaCl_2$ would depress the freezing point of water by only 0.006°C, which is far smaller than the scale of random temperature fluctuations within the freezer. Nevertheless, ice spikes were seven times less likely to occur in 0.001 molar calcium chloride than in pure water. The highest freezing-point depressions in the experiment, those for the 0.1 molar alcohol solutions, were only 0.186°C, a negligible amount. Moreover, the probability of ice-spike formation at a given concentration varied widely between solutes, suggesting that colligative properties are not at work. The data suggest that freezing-point depression could only play a significant role given two rather unlikely conditions. The first is that solutes reach concentrations orders of magnitude larger in drops at the tips of ice spikes than in the original solution. The second is that the scale of this amplification effect varies widely between solutes. A future experiment comparing concentrations of solutes at the tips of ice spikes with concentrations in the rest of the cube could investigate this model.

Since freezing-point depression is an unlikely source of ice-spike inhibition, a more likely model is based on crystal formation. According to the Bally-Dorsey model, ice spikes form when most of the surface of the water is frozen, leaving only a small hole. The convergence of three dendritic crystal sheets usually forms this triangular hole, which becomes the base of the ice spike. Inhibition of the formation of these sheets could easily account for reduced probabilities of ice-spike formation at large concentrations. Though this model seems promising, the dependence of crystal formation on the chemical properties of the substance is not well understood. An experiment directly observing crystal growth in water as a function of the concentration of chemical additives would be necessary to investigate this model.

Qualitative observations of ice lumps and transparent cubes present numerous avenues for future research. It is possible that the ice lumps observed at low concentrations are formed in a manner similar to ice spikes, whereby liquid water is pushed out through a hole in the nearly frozen surface. Lumps may be formed when water is pushed out too quickly, causing it to flow over the surface, or when the initial hole is too large, in which case the drop will freeze too slowly to lengthen into a spike. Since the appearance of ice spikes appears to be highly correlated with the appearance of ice lumps, it is likely that the same factors lead to the inhibition of both. The transparency of some cubes at high concentrations is a more perplexing phenomenon, as there is no readily apparent mechanism for the dissipation of dissolved gases. This effect was only observed in the presence of liquid solvents, perhaps indicating that dissolved solids and liquids may accumulate differently in the ice during freezing. The correlation between the transparency of ice cubes and the inhibition of ice spikes could not even be qualitatively established within the context of this experiment. A longer study might make more careful inspections of the cubes in addition to testing more solutes to gain a better understanding of this effect.

5 Conclusion

Chemical additives in water have been found to inhibit the formation of ice spikes. Metallic salts have been found to inhibit the formation of ice spikes to a large extent, even at low concentrations. Alcohols have been found to promote the formation of ice spikes, inhibiting ice-spike formation only at concentrations greater than or equal to 10^{-3} molar. This experiment has not supported a model of the inhibition of ice spikes based on freezing-point depression, but has suggested one based on the formation of dendritic sheets.

6 Acknowledgments

I would like to thank my mentor, Dr. Kenneth Libbrecht, physics professor and department chair at the California Institute of Technology, for his guidance and direction in completing this project. I would also like to thank my tutor, Dr. John Rickert, RSI technology assistant Jake Wildstrom, teaching assistants Emma Schmidgall and Bryant Mathews, and Molly Peeples for their aid in completing this research. I would like to especially express my gratitude to the Research Science Institute and the Center for Excellence in Education for making this research possible.

References

[1] G. Abrusci. Question #65. What conditions determine crystal growth? American Journal of Physics 65 (1997), no. 10, 941.

[2] O. Bally. Über eine eigenartige Eiskrystallbildung. Helvetica Chimica Acta 18 (1933), 475.

[3] D.C. Blanchard. Shorter contribution: a verification of the Bally-Dorsey theory of spicule formation on sleet pellets. Journal of Meteorology 8 (1951), 268–269.

[4] M. Chen. Ice spike formation induced by dendritic ice sheets. Undergraduate research report, University of Toronto (1993).

[5] H.G. Dorsey. Peculiar ice formations. American Physical Society 18 (1921), no. 2, 162–164.

[6] United States Environmental Protection Agency. Secondary drinking water regulations: guidance for nuisance chemicals. Available at `http://www.epa.gov/safewater/consumer/2ndstandards.html` (2004/08/01).

[7] United States Environmental Protection Agency. Ground water and drinking water: frequently asked questions. Available at `http://www.epa.gov/safewater/faq/faq.html` (2004/08/02).

[8] United States Environmental Protection Agency. List of contaminants and MCLs. Available at `http://www.epa.gov/safewater/mcl.html` (2004/08/02).

[9] L. Evans, R. McLachlan, and S. Morris. The spikes/whiskers phenomenon. The Physics Teacher 31 (1993), no. 5, 264–265.

[10] J. Hallett. Crystal growth and the formation of spikes in the surface of supercooled water. Journal of Glaciology 3 (1960), no. 28, 698–704.

[11] L. Hill, E. Lozowski, and R.D. Sampson. Experimental Evidence on Ice Spikes.

[12] V. Jamieson. Freezer teaser. New Scientist 180 (2003), 36.

[13] K. G. Libbrecht. Ice spikes. Available at `http://www.its.caltech.edu/~atomic/snowcrystals/icespikes/icespikes.htm` (2004/07/02).

[14] K.G. Libbrecht and K. Lui. An investigation of laboratory-grown ice spikes. To appear in Journal of Glaciology.

[15] S. Morris. Got spikes on your ice cubes? Available at `http://www.physics.utoronto.ca/~smorris/edl/icespikes/icespikes.html` (2004/07/02).

[16] D.L. Ozsvath. Groundwater quality. Available at `http://www.uwsp.edu/geo/faculty/ozsvath/lectures/Groundwater_Quality.htm` (2004/08/01).

[17] H.F. Perry. Ice spikes—can you explain them? The Physics Teacher 31 (1993), 112.

[18] H.F. Perry. A last word on ice spikes. The Physics Teacher 33 (1995), 148–149.

[19] H.F. Perry. Answer to Question #65. What conditions determine crystal growth? American Journal of Physics 69 (2001), no. 2, 106.

On a Generalization of the Collatz Conjecture

John Lesieutre

under the direction of
Mr. Zuoqin Wang
Massachusetts Institute of Technology

Abstract

In this paper we analyze a generalized version of the Collatz conjecture proposed by Zhang Zhongfu and Yang Shiming. We present a heuristic argument in favor of their conjecture and generalize a number of fundamental theorems from the original problem. We then obtain results related to properties of the generalized conjecture which do not arise in the original.

1 Introduction

The Collatz conjecture is one of the long-standing open problems of mathematics. Apparently first posed by Lothar Collatz in the 1930s, it has since withstood every attempt at proof. Since the problem was passed around orally for many years before any articles were written about it, it goes by a number of names, including Kakutani's problem, Hasse's algorithm, Ulam's problem, and the Syracuse problem. The problem is related to a wide range of topics in mathematics, including number theory, computability theory, and the analysis of dynamical systems. The statement of the Collatz conjecture involves the mapping $T : \mathbb{N} \to \mathbb{N}$ where

$$T(x) = \begin{cases} x/2 & \text{if } x \text{ is even,} \\ 3x+1 & \text{if } x \text{ is odd.} \end{cases}$$

Let $T^{(i)}(x)$ denote the result of i iterations of T on x. We call the sequence of $T^{(i)}(x)$ the *trajectory* or x. The Collatz conjecture states that for all x there is some i such that $T^{(i)}(x) = 1$. For example, if we start with $x = 7$, the iteration goes $7 \to 22 \to 11 \to 34 \to 17 \to 52 \to 26 \to 13 \to 40 \to 20 \to 10 \to 5 \to 16 \to 8 \to 4 \to 2 \to 1$. The conjecture has been verified for initial values of x up to 2^{50}, but proof remains elusive. However, there exist many partial results; one theorem shows that asymptotically, "nearly all" numbers do eventually reach 1, though there still may be an infinite number of exceptions. Other theorems estimate the number of iterations a typical number takes to reach 1 or show that the conjecture holds for certain broad classes of integers (see Lagarias's survey [5]).

Here we deal with a generalization of the Collatz conjecture proposed by Zhang Zhongfu and Yang Shiming [8]. We examine the mapping $T_n : \mathbb{N} \to \mathbb{N}$ defined by the rule

$$T_n(x) = \begin{cases} \dfrac{x}{p_{i_1} p_{i_2} \cdots p_{i_k}} & \text{where the } p_i \text{ are exactly the primes } \le p_n \text{ that divide the numerator,} \\ p_{n+1}x + 1 & \text{if no } p_i \le p_n \text{ divides } x. \end{cases}$$

For example, with $n = 3$, we multiply by 7, add 1 and then divide out by 2, 3, and 5. Taking $x = 37$ as our initial value, the trajectory of x under iteration of T_n is $37 \to 260 \to 26 \to 13 \to 92 \to 46 \to 23 \to 162 \to 27 \to 9 \to 3 \to 1$. Here we investigate the properties of this function under iteration. We note that our general case contains nontrivial cycles, which are conjectured not to occur in the original problem, and thus formulate our motivating conjecture:

Conjecture 1. *For any p_{n+1} and x_0, repeated iteration of $T_n(x)$ on x_0 will eventually yield $T_n^{(i)}(x_0) = 1$ or $T_n^{(i)}(x_0)$ will enter one of a finite number of nontrivial periodic cycles.*

These nontrivial cycles are discussed more fully in Section 3.4. A small example is the periodic trajectory $17 \to 47 \to 37 \to 17 \to \cdots$ for $p_{n+1} = 11$.

The Collatz conjecture is very similar to the case $n = 1$ in our more general statement, except that it also excludes the possibility of nontrivial cyclic trajectories. A proof of the general conjecture would not resolve the Collatz conjecture, as there could still exist nontrivial cycles. It is known [5] that if such cycles exist they must have period at least 275,000.

While a great body of work pertaining to the original conjecture (our $n = 1$) exists, the only paper concerning this generalization is the paper by Zhang Zhongfu and Yang Shiming in which the generalized form is suggested. They do not analyze in it much depth, only demonstrating the existence of nontrivial cycles and discussing a slight further generalization.

Using a wide range of techniques, mathematicians have proved numerous results about the Collatz conjecture, from demonstrating that it holds for "almost all" large x to showing the truth of certain cases to demonstrating its connections to rational approximations of $\log_2 3$. It has been attacked with graph theory, analysis, elementary number theory, and ergodic theory, with each of these approaches yielding new results. For a more complete overview, the reader may consult Lagarias's excellent survey article [5].

2 Probabilistic Analysis of Trajectories

Our conjecture is equivalent to the assertion that there are only a finite number of cycles and that the mapping contains no divergent trajectories, that is, there is no x_0 for which $T_n^{(i)}(x_0) \to \infty$ as $i \to \infty$. If we assume that for large initial values, the function acts sufficiently randomly that residues are equally distributed modulo $\left(p_n^{\#}\right)^k = (p_1 p_2 \ldots p_n)^k$, we may calculate expected trends in the size of the numbers in our trajectory and then supply some supporting evidence for part of our conjecture. We specifically investigate the expected ratio between consecutive terms not divisible by primes $\leq p_n$ in trajectories. For a large integer n_0, we may estimate the ratio between n_0 and n_1, the next term coprime to $p_n^{\#}$.

Heuristic 1. *We expect that*

$$\frac{n_1}{n_0} \approx p_{n+1} \prod_{j=1}^{n} p_j^{-p_j/(p_j-1)^2}.$$

Proof. Before proceeding, we must establish two preliminary lemmata which are necessary to our argument.

Lemma 2. *If set A is a reduced residue system modulo $p_n^{\#} = p_1 p_2 \ldots p_n$, then any p_i with $i \leq n$ divides precisely $\#(A)/(p_i - 1)$ elements of the set $kA+1$ for any k not divisible by any prime less than or equal to p_n.*

Proof. Consider the set of residues of A taken mod p_i. There are no 0 residues by definition of A, since $p_i \mid p_n^{\#}$. However, the residues must be evenly distributed over the other $p_i - 1$ possibilities. Suppose that some residue k_1 in A occurs more often than some other residue k_2. When we multiply A by a constant relatively prime to the modulus $p_n^{\#}$, we get the same residue system. But there is one such number C so that $k_1 \equiv C k_2 \pmod{p_n^{\#}}$. Therefore k_1 and k_2 must occur with equal frequency among the elements of A taken mod p_i. So the residue -1 occurs $\frac{\#(A)}{p_i - 1}$ times in the set A taken mod p_i. Now, the set kA is the same residue system as A for $(k, p_n^{\#}) = 1$. So the residue -1 occurs there this same number of times. Now, in the set $kA+1$ the residue 0 occurs exactly this many times, and exactly that many elements of $kA+1$ are divisible by p_i. □

Lemma 3. *We have*

$$\prod_{i=1}^{\infty} (n^{-i})^{(n^{-i})} = n^{-n/(n-1)^2}$$

for $n > 1$.

Proof. First, note that

$$\begin{aligned}
\prod_{i=1}^{\infty} (n^{-i})^{(n^{-i})} &= \exp\left(\log \prod_{i=1}^{\infty} (n^{-i})^{(n^{-i})}\right) \\
&= \exp\left(\sum_{i=1}^{\infty} n^{-i} \log n^{-i}\right) \\
&= \exp\left(-\log n \sum_{i=1}^{\infty} i n^{-i}\right).
\end{aligned}$$

So now let $k = \sum_{i=1}^{\infty} i n^{-i} = 1/n + 2/n^2 + 3/n^3 + \cdots$ so that

$$\begin{aligned}
nk &= 1 + \frac{2}{n} + \frac{3}{n^2} + \frac{4}{n^3} + \cdots \\
&= \left(1 + \frac{1}{n} + \frac{1}{n^2} + \cdots\right) + \left(\frac{1}{n} + \frac{2}{n^2} + \frac{3}{n^3} + \cdots\right) \\
&= \frac{1}{1 - \frac{1}{n}} + k,
\end{aligned}$$

whence $k = n/(n-1)^2$. Thus our product is equal to

$$\begin{aligned}
\exp\left(-\log n \sum_{i=1}^{\infty} i n^{-i}\right) &= \left(e^{\log n}\right)^{-n/(n-1)^2} \\
&= n^{-n/(n-1)^2},
\end{aligned}$$

as desired. □

Now we return to the heuristic. The first step is to determine the expected factor by which a given prime p_i will decrease a number when we divide out by small factors. Consider a prime $p_i \leq p_n$. The probability that p_i will divide $p_{n+1}x+1$ exactly m times is equal to the probability that it will divide it once, times the probability that it will divide it $m-1$ more times, times the probability that it will divide it no more times after that. By Lemma 2 with $k = p_{n+1}$ and $n_0 \in A$, the chance that $p_{n+1}n_0+1$ is divisible by p_i once is $1/(p_i-1)$. The chance that $p_{n+1}n_0+1$ will be divisible by p_i exactly m times is then

$$\left(\frac{1}{p_i-1}\right)\left(\frac{1}{p_i}\right)^{m-1}\left(\frac{p_i-1}{p_i}\right) = \frac{1}{p_i^m},$$

as described above. It may seem irrational that we treat the first division differently than the others, but to see why this is plausible, consider the $3x+1$ mapping of the Collatz conjecture; when we apply $3x+1$ to an odd number we may always take out a factor of 2 immediately, but after that we treat it as random. Now, when p_i divides a number m times, the factor of reduction is $1/p_i^m$. We then expect p_i to reduce $p_{n+1}x+1$ by an average factor of

$$\prod_{m=1}^{\infty} (1/p_i^m)^{\left(\frac{1}{p_i^m}\right)} = \prod_{m=1}^{\infty} (p_i^{-m})^{(p_i^{-m})} = p_i^{-p_i/(p_i-1)^2}$$

by Lemma 2. The expected total factor of reduction due to all primes up to p_n is therefore

$$\prod_{j=1}^{n} p_j^{-p_j/(p_j-1)^2}.$$

The expected ratio is then found as

$$\frac{n_1}{n_0} \approx p_{n+1}\prod_{j=1}^{n} p_j^{-p_j/(p_j-1)^2}.$$

$\square$

For $n=1$, the Collatz conjecture, the expected ratio is 3/4, which suggests that on a large scale, trajectories tend to decrease and divergence is unlikely. For $n=2$, the expected ratio is $\frac{5\sqrt[4]{3}}{12} \approx 0.54836$, which suggests that trajectories will tend to decrease even faster. We might expect that this ratio would tend to 0 as n increases without bound since we are dividing out by an ever-larger number of primes. We claim that instead this ratio converges: note first that

$$\begin{aligned}
&p_{n+1}\prod_{j=1}^{n} p_j^{-\frac{p_j}{(p_j-1)^2}} \\
&= \exp\left(\log p_{n+1}\sum_{j=1}^{n}\log p_j^{-p_j/(p_j-1)^2}\right) \\
&= \exp\left(\log p_{n+1} - \sum_{j=1}^{n}\frac{p_j}{(p_j-1)^2}\log p_j\right) \\
&= \exp\left(\log p_{n+1} - \sum_{j=1}^{n}\left(\frac{1}{p_j} + \frac{1-2p_j}{p_j(p_j-1)^2}\right)\log p_j\right) \\
&= \exp\left(\log p_{n+1} - \sum_{j=1}^{n}\frac{\log p_j}{p_j} - \sum_{j=1}^{n}\frac{(1-2p_j)\log p_j}{p_j(p_j-1)^2}\right).
\end{aligned}$$

It is a short exercise in analytic number theory to show that $\sum_{p\leq N}\frac{\log p}{p} = \log N + C_1 + o(1)$ for some constant C_1. Some simple calculus shows that the second sum converges over all the naturals, and therefore its sum over the primes is some constant C_2. Taking $N = p_{n+1}$ above, we find that our product is

$$e^{\log p_{n+1} - \log p_{n+1} - C_1 - C_2 - o(1)} = e^{-(C_1+C_2+o(1))},$$

which converges to some constant C as $n \to \infty$. Calculating the value over the first 10,000 primes suggests that $C \approx 0.520$. We also note that the expected ratio seems to be less than 1 for all p_{n+1}, and our apparently tight error bound combined with our calculations over the first 10,000 p_{n+1} provides strong evidence for this.

3 Structure of Trajectories

It is very difficult to directly analyze the trajectory of a given number, but we may still obtain some results. To begin, we present generalizations of several basic results of Terras and others presented in Lagarias's paper [5].

3.1 The Encoding Matrix

One of the most important tools in the study of the Collatz conjecture has been Terras's encoding vector (called the parity vector by Lagarias), which contains entries corresponding to the action of the function for each iteration (whether the image of x is $3x+1$ or $x/2$) [7]. In our general case we define $A_k(n)$ to be the matrix whose entries m_{ij} satisfy $0 \leq m_{ij} < p_i$ and $m_{ij} \equiv T_n^{(k)}(x) \pmod{p_i}$. The matrix represents the action of the function on the initial number under repeated iteration. For example, the iteration of T on 13 with $p_{n+1} = 7$ yields $13 \to 92 \to 46 \to 23 \to 162 \to 27 \to 9 \to 3 \to 1$. As an encoding matrix, we write

$$A_9(13) = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 2 & 0 & 0 & 0 & 0 & 1 \\ 3 & 2 & 1 & 3 & 2 & 2 & 4 & 3 & 1 \end{pmatrix}.$$

For example, the 3rd column corresponds to 46, which is congruent to 0 mod 2, 1 mod 3, and 1 mod 5. The final column represents 1. We denote the n^{th} column by $\alpha_n(x)$ and the vector of the n^{th} row as $\beta_n(x)$. We may use the encoding matrix to study the trajectory of x. We may derive an important formula for $T_n^{(k)}(x)$ in terms of $A_k(x)$.

Explicitly, letting α_i denote the i^{th} column of A_k, β_i the i^{th} row of A_k and defining

$$\begin{aligned} a(A_k, j) &= \#\{\alpha_i : j \le i \le k, \alpha_i \text{ has no zero entry}\}, \\ b(\beta_i, j) &= \#\{m_{ie} : j \le e \le k, m_{ie} = 0\}, \end{aligned}$$

and

$$c(\alpha_i) = \begin{cases} 0 & \text{if some entry of } \alpha_i \text{ is } 0, \\ 1 & \text{otherwise,} \end{cases}$$

we can decompose $T_n^{(k)}(x)$ as follows.

Theorem 4. *We have* $T_n^{(k)}(x) = \lambda_k(x)x + \rho_k(x)$, *where*

$$\lambda_k(x) = \frac{p_{n+1}^{a(A_k,1)}}{\prod_{j \le n} p_j^{b(\beta_j,1)}}$$

and

$$\rho_k(x) = \sum_{i=1}^{k} \frac{c(\alpha_i) p_{n+1}^{a(A_k,i+1)}}{\prod_{j \le n} p_j^{b(\beta_j,i)}}.$$

Essentially, the quantity λ_k measures the overall multiplicative factor, and ρ_k makes small corrections for the +1s. Note that we may sometimes write $\lambda_i(A_k)$ and $\rho_i(A_k)$, and here A_k is the matrix representing the trajectory of x.

Proof. We proceed by induction on k. First, we handle the base case $k = 1$. If α_1 contains all nonzero entries, so that $a(A_1, 1) = 1$ and x is divisible by no prime $\le p_n$, the image of x is $T_n^{(1)}(x) = p_{n+1}x + 1$. Our formula gives

$$\lambda_1(x) = \frac{p_{n+1}^1}{1} = p_{n+1}$$

and

$$\rho_1(x) = \sum_{i=1}^{1} \frac{p_{n+1}^{a(A_k,i+1)}}{\prod_{j \le n} p_j^{b(\beta_i,j)}} = \frac{p_{n+1}^0}{\prod_{j \le n} p_j^0} = 1.$$

Then we have $T^{(1)}(x) = \lambda_1(x)x + \rho_1(x)$ in this case. Now consider the case when α_1 has some entry equal to 0. In this case λ_1 has a numerator of 1 since $a(A_1, 1) = 0$ and a denominator equal to the product of all primes $\le p_n$ dividing x, while ρ_1 is 0 since $c(\alpha_1) = 0$. Then $T_n^{(1)}(x)$ is x divided by these primes, as it should be. So our base case holds.

So now assume that we have $T_n^{(k)}(x) = \lambda_k(x)x + \rho_k(x)$ for all k. When α_{k+1} has all nonzero entries, we should have $T_n^{(k+1)} = p_{n+1}T_n^{(k)} + 1$. We have $\lambda_{k+1} = p_{n+1}\lambda_k$ because $a(A_{k+1}, 1) - a(A_k, 1) = 1$ and the denominator remains unchanged. We also get

$$\rho_{k+1} = \sum_{i=1}^{k+1} \frac{c(\alpha_i) p_{n+1}^{a(A_k,i+1)}}{\prod_{j \le n} p_j^{b(\beta_i,j)}} = p_{n+1}\rho_k + 1$$

from the definitions of a, b, and c. So $T_n^{(k+1)}(x) = \lambda_{k+1}(x)x + \rho_{k+1}(x) = p_{n+1}\lambda_k(x)x + p_{n+1}\rho_k(x) + 1 = p_{n+1}\big(\lambda_k(x)x + \rho_k(x)\big) + 1 = p_{n+1}T_n^{(k)} + 1$, as it should. Now consider the case when α_{k+1} has entries equal to 0. Here $\lambda_{k+1}(x)$ is just $\lambda_k(x)$ divided by the primes dividing the most recent iterate $T_n^{(k)}(x)$. Additionally $\rho_{k+1}(x)$ is $\rho_k(x)$ divided by these primes, since $c_{k+1} = 0$ and our bottom changes by this factor. So we have $T_n^{(k+1)}(x) = T_n^{(k)}(x)$ divided out by small primes. This completes our proof. ☐

We may prove various other theorems about the densities and structures of trajectories, but first, we require some definitions.

3.2 Matrices: Attainable, Admissible, and Inflating

We call a matrix B_k *attainable* if there exists some n_0 such that $A_k(n_0) = B_k$.

To see why not all matrices are attainable, note that no encoding matrix may contain the sequence

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

since this would correspond to taking an odd number, dividing it by 3, and getting an even number. If two vectors α_1 and α_2 may occur as consecutive α_i and

α_{i+1}, we say that α_2 is a *successor* of α_1, and we denote the number of possible successors of a vector α by $n(\alpha)$. A matrix is attainable if and only if α_{i+1} is a successor of α_i for all $0 \leq i < k$.

We may further classify the attainable matrices. Following Lagarias, we say that a matrix A_k is *admissible with length k* if $\lambda_k(A_k) < 1$, while $\lambda_i(A_k) > 1$ for $1 \leq i < k$. Essentially, this means that the total multiplicative factor is greater than 1 for the first $k-1$ columns and less than 1 after k columns. Note that λ_k is never equal to 1, since it is a power of p_{n+1} over a product of powers of smaller primes. Similarly, we call a matrix A_k *inflating with length m* if $\lambda_i(A_k) > 1$ for $1 \leq i \leq m$. This means that the multiplicative factor is always greater than 1 for the first m columns of the matrix.

It is clear that there are $(p_n^{\#})^k$ different matrices of length k for a given p_n. But how many of these are attainable? It is difficult to analyze this in the general case but not altogether too difficult to solve the case $n = 2$ ($p_{n+1} = 5$). We find that the number of attainable matrices grows exponentially with k.

Theorem 5. *The number of attainable matrices of length k in the case $n = 2$ is asymptotically proportional to λ_1^k, where λ_1 is the unique real solution to $x^5 - x^4 - 2x^3 - 2x^2 - x - 1 = 0$.*

While this theorem does not apply to the general case, it suggests how it may be solved and also that the frequency of allowed matrices should decrease as k increases.

Proof. We have noted that a matrix is admissible if and only if each α is a successor of its predecessor. Therefore, in estimating the number of attainable matrices, we need only consider the last α. We can then easily derive relations between the number of attainable matrices for k and $k+1$. To begin, we consider all possible pairs of successors. These are given in Figure 1.

Denoting by $a_k, b_k, \ldots, f_k$ the number of $\binom{0}{0}$'s, $\binom{0}{1}$'s, $\ldots$, and $\binom{1}{2}$'s ending attainable matrices of length k, we derive the system

$$\begin{aligned}
a_{k+1} &= a_k + e_k \\
b_{k+1} &= a_k + c_k \\
c_{k+1} &= a_k + b_k + f_k \\
d_{k+1} &= a_k + d_k \\
e_{k+1} &= a_k + c_k + d_k \\
f_{k+1} &= a_k + b_k + d_k
\end{aligned}$$

$$\begin{aligned}
&\begin{pmatrix} 0 \\ 0 \end{pmatrix} \to \begin{pmatrix} 0 \\ 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix}\begin{pmatrix} 0 \\ 2 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \\
&\begin{pmatrix} 0 \\ 1 \end{pmatrix} \to \begin{pmatrix} 0 \\ 2 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \\
&\begin{pmatrix} 0 \\ 2 \end{pmatrix} \to \begin{pmatrix} 0 \\ 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \\
&\begin{pmatrix} 1 \\ 0 \end{pmatrix} \to \begin{pmatrix} 1 \\ 0 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \\
&\begin{pmatrix} 1 \\ 1 \end{pmatrix} \to \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \\
&\begin{pmatrix} 1 \\ 2 \end{pmatrix} \to \begin{pmatrix} 0 \\ 2 \end{pmatrix}.
\end{aligned}$$

Figure 1: Possible successors

for $k \geq 1$, with the initial condition $a_1 = \cdots = f_1 = 1$. These Fibonacci-type recurrences guarantee that our sum $a_i + b_i + c_i + d_i + e_i + f_i$ grows essentially exponentially. To see this, we rewrite the system as $X_{k+1} = AX_k$, where X_k is the column vector $(a_k, b_k, \ldots, f_k)$ and A is the matrix

$$\begin{pmatrix}
1 & & & & 1 & \\
1 & & 1 & & & \\
1 & 1 & & & & 1 \\
1 & & & 1 & & \\
1 & & 1 & 1 & & \\
1 & 1 & & 1 & &
\end{pmatrix}.$$

Then $X_k = A^{k-1}X_1$ for $k \geq 1$, where X_1 is the column vector $(1, 1, 1, 1, 1, 1)$. By machine computation one finds that the eigenvalues of A are 1 and the roots of $\lambda^5 - \lambda^4 - 2\lambda^3 - 2\lambda^2 - \lambda - 1 = 0$, of which $\lambda_1 \approx 2.335$ is the largest in absolute value. Furthermore, the coefficient in X_1 of the eigenvector corresponding to λ_1 is nonzero. Consequently the sum of the entries of X_k, namely $a_k + \cdots + f_k$, is asymptotic to a constant times λ^k.

$\square$

In particular, because $\lambda_1 < 6$, allowed matrices become increasingly less frequent (as a fraction of all matrices) as the length k increases.

3.3 Some Fundamental Results on Encoding Matrices

The encoding matrix possesses a number of interesting properties that make it useful in the study of our problem. We generalize two important theorems of Terras below [7].

Our first result states that the numbers whose encoding matrices of length k are identical fall in arithmetic sequences.

Theorem 6. *Let* $S(A_k) = \{x : A_k(x) = A_k\}$ *for an attainable matrix* A_k*. Then* $S(A_k) = \{x_0(A_k) + ip_n^{\#}\prod_{j=1}^{k-1} n(\alpha_j) : i \geq 0\}$*, where* $n(\alpha_j)$ *is the number of possible successors to* α_j*, and* $x_0(A_k)$ *is the least* x *satisfying* $A_k(x) = A_k$*.*

Proof. We proceed by induction on k, the length of A. Clearly, when $k = 1$ our theorem holds. By the Chinese Remainder Theorem, some x_0 exists in the interval $[1, p_n^{\#}]$ satisfying $A_1(x_0) = A_1$. Also, it is clear that $A_1(x_0) = A_1(x)$ if and only if $x = x_0 + p_n^{\#} i$. Further, we see that after we take the image of the members of this sequence under A_1, they must be evenly distributed over the possible successors of α_1 when taken modulo $p_n^{\#}$. To see this, consider two cases: when we divide by some product of primes $\leq p_n$, we get a sequence of numbers with some fixed residues modulo each p_j by which we did not divide and with all possible choices for the residue modulo each p_i by which we did divide, and these are equally distributed. The number of these possibilities is exactly $n(\alpha_1)$. And when we map by $p_{n+1}x + 1$, we have only one possible residue modulo $p_n^{\#}$, which is still the equidistribution we seek.

Now assume the theorem holds for a length k. That is, we assume $S(A_k) = \{x_0(A_k) + ip_n^{\#}\prod_{j=1}^{k-1} n(\alpha_j) : i \geq 0\}$. After we take the image of some x under the mapping described by A_k, we must be left with one of the $n(\alpha_k)$ possible residues. None of these possible residues occurs more often than any other by an argument analogous to that above. Our sequence must therefore be split in $n(\alpha_k)$ new arithmetic sequences with the same common difference. Therefore the common difference is multiplied by this factor $n(\alpha_k)$. There must also still exist some x_0 for each A_k. These facts imply our inductive step and thereby our theorem. □

We generalize two functions of Terras [7]. Let $\sigma(x)$, the *stopping time* function, be the smallest k so that $T_n^{(k)}(x) < x$, that is, the number of iterations it takes for a number to reach a number less than itself. In the Collatz problem, a proof that every number has a finite stopping time would amount to a proof of the conjecture since we could inductively show that each number > 1 must reach a smaller number, which in turn must reach 1 by the induction hypothesis. Also let $\omega(x)$, the *coefficient stopping time* function, be the smallest k such that $\lambda_k(x) < 1$. Then let $H_k = \{x : \sigma(x) = k\}$ and $I_k = \{x : \omega(x) = k\}$. It is clear that $H_k \subseteq I_k$. In fact the inclusion is not far from an equality.

Theorem 7. $I_k \setminus H_k$ *is a finite set.*

Proof. Our proof closely follows that of Terras' for the Collatz problem [7]. To begin, note that when A_k is an admissible matrix of length k and $x \in S(A)$, then for $1 \leq i \leq k-1$ we have

$$T_n^{(i)}(x) \geq \lambda_k(A_k)x \geq x$$

by the definition of admissibility. Thus, all x with length-k matrix A_k have a stopping time $\sigma(x) \geq k$. Now define $\epsilon_k(A_k) = 1 - \lambda_k(A_k)$ for admissible A_k. Then we have $T_n^{(k)}(x) = x + \big(\rho_k(A_k) - \epsilon_k(A_k)x\big)$. The number x has stopping time exactly k whenever $\rho_k(A_k) - \epsilon_k(A_k)x < 0$, that is, whenever $x > \rho_k(A_k)/\epsilon_k(A_k)$. Since this depends only on A_k, there are only finitely many $x \in S(A_k)$ for which $\sigma(x) \neq \omega(x)$. So if A_k is admissible then all sufficiently large x in $S(A_k)$ have stopping time k.

Now suppose that A_k is not admissible. If no initial segment of A_k is admissible, then A_k is inflating and $\lambda_k(A_k) > 1$, so no members of $S(A_k)$ have stopping time k. If some initial segment of A_k is admissible, then both $\omega(A_k)$ and $\sigma(A_k)$ are less than k for sufficiently large elements of $S(A_k)$, and only a finite number of elements of $S(A_k)$ may have stopping time k. This, taken with our first argument and the observation that there are only a finite number of matrices of length k, proves our theorem. □

Corollary 8. H_k *and* I_k *have the same asymptotic density. Further, this asymptotic density is given by*

$$F(k) = \sum_{\substack{A \text{ admissible} \\ \text{length}(A) \leq k}} \text{weight}(A),$$

where

$$\text{weight}(A) = \left(p_n^{\#}\prod_{j=1}^{k-1} n(\alpha_j)\right)^{-1}.$$

Proof. Both H_k and I_k are infinite sets. They share all but a finite number of elements, which implies that they have equal asymptotic density. Our weight function is derived from Theorem 6 and makes sense on a fairly intuitive level; when all possible successor matrices to some matrix are still admissible, our weight remains unchanged. But when some of these are not, our weight corresponding to that particular branch is reduced by some factor. The sum of the weights over all admissible vectors of length k must be the density of numbers with stopping time $\leq k$. $\square$

Conjecture 2. $\lim_{k\to\infty} F(k) = 1$.

We conjecture that the density of numbers with finite stopping time is 1. This is a known result in the Collatz conjecture. We have attempted to resolve this conjecture by using combinatorial methods to bound the sizes of the sets of attainable, admissible, and inflating vectors of length k and then applying a bound on the weight function. However, our methods have not achieved strict enough bounds to ensure convergence to 1.

3.4 Nontrivial Cycles

One conjecture about the Collatz mapping is that for our $n = 1$, there are no nontrivial cycles (i.e., no numbers go into repeating loops that do not include 1). However, in our general case, many such cycles exist and we may find them fairly easily. As noted in the introduction, the simplest such cycle occurs with $p_{n+1} = 11$, when we have the loop $17 \to 47 \to 37 \to 17 \to \cdots$. A given p_{n+1} may lead to multiple such cycles; $p_{n+1} = 61$ seems to be the smallest such case, though we have not proven this. We may show that a given p_{n+1} cannot have too many small nontrivial cycles:

Theorem 9. *For a given p_{n+1}, there exist only finitely many cycles of any given length $k \geq 1$.*

Proof. Our theorem is a consequence of Theorem 7. Consider the set C of smallest elements of cycles of length k. All elements of C have some finite $\omega(x)$ since $T_n^{(k)}(x) = x$, so that $\lambda_k(A_k(x)) < 1$ because $\rho_k(A_k(x))$ must be positive. However, no $x \in C$ has a finite $\sigma(x)$, since x is the smallest element of the cycle. By Theorem 7, the set C is finite. $\square$

Of course, we expect that there are only finitely many cycles of *any* length, but it is not clear how to prove this. Nor do we know exactly how many cycles we expect for a given p_{n+1}, or even whether large p_{n+1} contain more or less cycles than smaller ones. The prime $p_{n+1} = 61$ seems to be the only prime less than 300 for which more than one cycle exists, but we cannot be sure that there are other cycles that eluded our computer search.

4 A Further Generalization

We may consider a more general version of our problem, replacing $p_{n+1}x + 1$ in the definition of $T_n(x)$ with $p_{n+1}x + K$ for some K with $(K, p_n^{\#}) = 1$. We still divide out by primes $\leq p_n$ as in our original problem. These generalization is also posed and briefly discussed by Zhang and Yang [8]. We may derive several simple theorems based on this map relating cycles between two different K in certain cases.

4.1 Basic Results

Theorem 10. *A cycle in the mapping using $p_{n+1}x + k$ may be transformed to one in $p_{n+1}x + kq$ when q is not divisible by any prime $\leq p_n$.*

Proof. Consider a trajectory $m_1 \to m_2 \to \cdots \to m_i \to \cdots$ that eventually becomes cyclic and is in its cycle for $i \geq s$. We have $m_{s+1} = (p_{n+1}m_s + k)/F_{s+1}$ where the F_i are divisible only by primes $\leq p_n$. So now $(qm_{s+1}) = \big(p_{n+1}(qm_s) + kq\big)/F_i$ and we see that $(qm_s, qm_{s+1}, \ldots)$ is a cycle in the mapping with $p_{n+1}x + kq$. $\square$

We may also obtain the following stronger but less general result:

Theorem 11. *Suppose we replace $p_{n+1}x + 1$ by $p_{n+1}x + p_{n+1}^a k$ in the definition of $T_n(x)$. Then for a given k and p_{n+1}, all mappings of this form contain the same number of cycles.*

Proof. As above, consider a trajectory $m_0 \to m_1 \to m_2 \to \cdots \to m_i \to \cdots$ so that $p_{n+1}m_{i-1} + p_{n+1}^a k = F_i m_i$ with F_i containing only prime factors $\leq p_n$ and m_i containing no such factors. Now define l_i and n_i so that $m_i = p_{n+1}^{l_i} n_i$ with $p_{n+1} \nmid n_i$, and consider two cases. Suppose first that $l_i \geq a$. Then

$$\begin{aligned} m_{i+1} &= (p_{n+1}m_i + p_{n+1}^a k)/F_{i+1} \\ &= \big(p_{n+1}(p_{n+1}^{l_i} n_i) + p_{n+1}^a k\big)/F_{i+1}; \\ p_{n+1}^{l_{i+1}} n_{i+1} &= p_{n+1}^a (p_{n+1}^{l_i - a + 1} + k)/F_{i+1}, \end{aligned}$$

and we have $l_{i+1} = a$. If instead $l_i < a$, then

$$\begin{aligned} m_{i+1} &= (p_{n+1}m_i + p_{n+1}^a k)/F_{i+1} \\ &= \big(p_{n+1}(p_{n+1}^{l_i} n_i) + p_{n+1}^a k\big)/F_{i+1}; \\ p_{n+1}^{l_{i+1}} n_{i+1} &= p_{n+1}^{l_i + 1} (n_i + p_{n+1}^{a - l_i - 1} k)/F_{i+1}, \end{aligned}$$

and so $l_{i+1} = l_i + 1$. Combining these two cases, we see that there is some t such that $p_{n+1}^a \mid m_j$ and $p_{n+1}^{a+1} \nmid m_j$ for all $j \geq t$. Now note that we have

$$m_{t+1} = (p_{n+1}m_t + p_{n+1}^a k)/F_{t+1},$$

or

$$\frac{m_{t+1}}{p_{n+1}^a} = \left(p_{n+1}\frac{m_t}{p_{n+1}^a} + k\right)/F_{t+1},$$

so our cycle in the $p_{n+1}x + p_{n+1}^a k$ gives us a cycle in $p_{n+1}x + k$, which is the case $a = 0$. By the preceding theorem, taking $q = p_{n+1}$, we may use induction and see that our cycle in $p_{n+1}x + p_{n+1}^{a_1}k$ corresponds to exactly one cycle in $p_{n+1}x + p_{n+1}^{a_2}k$ for any choice of a_1 and a_2 and both of these mappings contain exactly the same number of cycles. □

4.2 Prime 2-Cycles in the Case $n = 1$

Looking at examples, we notice that the $3x+K$ problem sometimes contains cycles of length 2, with $x_1 \to x_2 \to x_1 \to \cdots$, and in some of these $(x_1, x_2) = 1$. We call such cycles with coprime x_1, x_2 and length 2 *prime 2-cycles*. For example, with $K = 7$, we have $5 \to 11 \to 5 \to \cdots$. We completely characterize these cycles in the case $n = 1$.

Theorem 12. *All prime 2-cycles in the* $3x + K$ *problem are of the form* $x_1 = \frac{2^a+3}{(2^a+3,2^b+3)}$, $x_2 = \frac{2^b+3}{(2^a+3,2^b+3)}$, *and* $K = \frac{2^{a+b}-9}{(2^a+3,2^b+3)}$.

Proof. We seek solutions to the equations $\frac{3x_1+K}{2^a} = x_2$ and $\frac{3x_2+K}{2^b} = x_1$. Eliminating x_2 from this system yields

$$(2^{a+b} - 9)x_1 = (2^a + 3)K,$$

from which we see that x_1 must be some constant multiple of $(2^a + 3)/(2^a + 3, 2^{a+b} - 9)$. But for our cycle to be prime, it must be equal to exactly that or x_1 and K have a common factor. Also note that $(2^a + 3, 2^{a+b} - 9) = (2^b(2^a+3), 2^{a+b} - 9)$ since 2^b is coprime to both these odd numbers. Now $(2^a+3, 2^b(2^a+3) - (2^{a+b} - 9)) = (2^a+3, 3(2^b+3)) = (2^a+3, 2^b+3)$. Then we have $x_1 = (2^a + 3)/(2^a + 3, 2^b + 3)$ and we may substitute to obtain $K = (2^{a+b} - 9)/(2^a + 3, 2^b + 3)$ and $x_2 = (2^b+3)/(2^a+3, 2^b+3)$. In substituting, we also note that 2^a and 2^b are in fact the largest powers of two by which we may divide. Thus all solutions are of our form and all numbers of our form are actual solutions. □

4.3 Generalization of Other Results

We may also generalize most of our earlier results to this case, though we omit these proofs. A simple extension of Lemma 2 allows us to use the same heuristic and arrive at the same argument against the existence of divergent trajectories. Our formula for $\lambda_k(x)$ remains unchanged (since it is simply a multiplicative factor and does not consider the added terms), and we change $\rho_k(x)$ so that

$$\rho_k(x) = K\sum_{i=1}^{k} \frac{c(\alpha_i)p_{n+1}^{a(A_k,i+1)}}{\prod_{j \leq n} p_j^{b(\beta_j,i)}}.$$

This result is clearly analogous to our original formula for $\rho_k(x)$ except it is now multiplied by a constant factor of K. Using this result and others, we may use proofs essentially identical to those above to show that there are a finite number of cycles of each length for $p_{n+1}x + K$. One interesting difference is that since we require $x > \rho_k(A_k)/\epsilon_k(A_k)$ and ρ_k has been multiplied by a factor of K, we may have a larger set of x for which $\omega(x) \neq \sigma(x)$ and therefore possibly more cycles. Theorem 10 implies that $K = pq$ must contain at least as many cycles as $K = p$, which provides further evidence that the larger K, the more cycles we should expect. However, it remains difficult to analyze the predicted number of cycles for a given K.

5 Conclusion

We have analyzed a generalization of the Collatz conjecture and extended various fundamental results in the original problem to our version. We have given a heuristic argument against the existence of divergent trajectories. Further, we have shown that only a finite number of cycles may exist for a given p_{n+1} and have extended this result to a further generalization with $p_{n+1}x + K$. However, many questions remain. It is known that in the Collatz mapping almost all integers eventually iterate to 1. We desire a proof of an analogous result in the general case. Many problems also remain in the analysis of non-trivial cycles, especially in estimating their number for a given p_{n+1} and K. Of course, the greatest open problem is the generalized conjecture itself, a resolution of which would provide a strong new result on the Collatz conjecture and improve our understanding of such dynamic systems.

6 Acknowledgements

I am most grateful to my mentor, Mr. Zuoqin Wang of MIT, for his constant guidance, encouragement, and assistance. He located a problem for me to work on, found resources to read, suggested lines of research, and helped me along when I encountered difficulties. I am also indebted to my tutor, Chris Mihelich, for his helpful suggestions for improvements and careful editing, as well as for his mathematical help, especially with the lemma of analytic number theory used in section 2. I am also obliged to Kartik Venkatram, whose advice was invaluable during the revision process. I thank the Center for Excellence in Education, the Research Science Institute, and the RSI staff for making it possible for me to pursue this research. Finally, I thank my fellow RSI students for their support and for showing (or at least feigning) interest in my research.

References

[1] P.J. Andaloro. The $3x + 1$ problem and directed graphs. Fibonacci Quarterly 40 (2002), no. 1, 43–54.

[2] S. Andrei, M. Kudlek, and R. Niculescu. Some results of the Collatz problem. Acta Informatica 37 (2000), 145–160.

[3] R.B. Banerji. The $3x + 1$ problem and number representation. Available at `http://www.sju.edu/~rbanerji/rb/papers/paper1.htm` (2004/06/30).

[4] B. Brent. $3x + 1$ dynamics on rationals with fixed denominator. Available at `arXiv:math.DS/0203170` (2004/07/21).

[5] J.C. Lagarias. The $3x + 1$ problem and its generalizations. Available at `http://www.cecm.sfu.ca/organics/papers/lagarias/` (2004/06/29).

[6] K.R. Matthews. The generalized $3x + 1$ mapping. Available at `http://www.numbertheory.org/pdfs/survey.pdf` (2004/06/28).

[7] R. Terras. A stopping time problem on the positive integers. Acta Arithmetica 30 (1976), 241–252.

[8] Zhang Zhongfu and Yang Shiming. Ying She Shu Lie Wen Ti. Shu Xue Chuan Bo 22 (1998), no. 2, 76–88.

Evaluation of *cis-cis* HOONO Production by Laser Photolysis through Modeling of Reaction Kinetics

Da Lin

under the direction of
Dr. Mitchio Okumura and Mr. Andrew Mollner
California Institute of Technology

Abstract

Detection and study of *cis-cis* HOONO under atmospherically relevant conditions are expected to significantly influence current atmospheric models. Recent proposals have been made to utilize laser photolysis to produce *cis-cis* HOONO for spectroscopic detection under atmospheric pressures. A model was created using Facsimile to evaluate the feasibility of and determine the initial conditions necessary for experimental detection of *cis-cis* HOONO at 30, 100, and 500 torr of atmospheric pressure and 298 K. Studies identified sixteen reactions that are influential in NO_x activity in an experimental reaction cell. Simulations based on ordinary differential equations show that the percent *cis-cis* HOONO yield should be 5.27%, 7.57%, and 14.1% respectively, which are in good agreement with past theoretical and experimental data. Analysis of predicted models identifies the importance of several free-radical concentrations for experimental studies.

1 Introduction

During the past century, atmospheric pollution has become a growing concern in our society. An increase in tropospheric ozone has been observed globally, and the potential effects of this increase are far-ranging. High ozone levels will disrupt photosynthesis and growth in plants. Ozone has also been linked to respiration problems, tissue decay, the promotion of scar-tissue formation, and cell damage in humans [1].

The three-body association reaction

$$OH + NO_2 + M \longrightarrow HONO_2 + M \tag{1}$$

is among the most important processes of the oxidation chemistry of the Earth's lower atmosphere. Free radicals in both the HO_x and NO_x families play central roles in the destruction of ozone in the stratosphere, as well as in chain reactions that produce photochemical smog in the troposphere (see Figure 1) [2]. Reaction (1) mediates the free-radical concentration in the atmosphere through the conversion of reactive HO_x and NO_x species into the chemically inactive and photochemically stable species nitric acid (HNO_3). Thus, detailed knowledge of the rate and mechanism for reaction (1) is critical for accurate modeling of atmospheric chemistry.

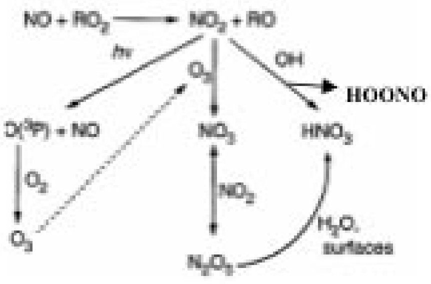


Figure 1: Summary of major reaction paths for NO_x in the lower atmosphere [9].

An intensive study of the large discrepancy between OH loss measurements in low and high pressures of bath gas (M) has led to several proposals of a reversible secondary channel

$$OH + NO_2 + M \longleftrightarrow HOONO + M \tag{2}$$

that produces HOONO (peroxynitrous acid), a more reactive isomer of $HONO_2$ [3]. Recent spectroscopic

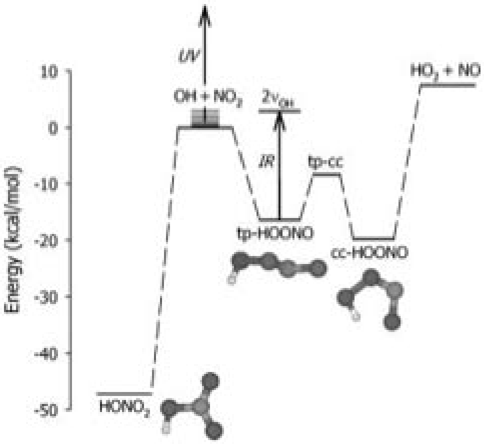


Figure 2: Energy-level diagram for the conversion of $OH + NO_2$ into $HONO_2$, HOONO, and $HO_2 + NO$ products. The *cc* and *tp* conformers of HOONO are shown. Binding energies of the *cc* and *tp* conformers and the isomerization barrier between them are taken from *ab initio* calculations [5].

experiments [4–6] have confirmed the existence of HOONO as a secondary product. An energy-level diagram of the two channels is shown in Figure 2. *Ab initio* calculations predict that the HOONO isomer is less stable than $HONO_2$ by 29.0 kcal/mol and that HOONO can exist in two different conformations, *cis-cis* (*cc*) and *trans-perp* (*tp*), where the labels refer to the ONOO and NOOH torsional angles [5]. The *cc*-HOONO conformer (19.8 kcal/mol) is reported to be more stable than the *tp*-HOONO conformer by 3.4 kcal/mol, with a significant isomerization barrier of 11.4 kcal/mol separating the two species [5].

Current estimates of the HOONO yield vary from less than 5% to more than 20% of the $HONO_2$ primary product [5]. A significant production of HOONO under atmospheric temperature and pressure would greatly affect current models of NO_x chemistry and ozone budgeting in the lower atmosphere. Formation of HOONO would reduce the efficiency of reaction (1) as a sink of OH and NO_2 radicals and notably affect the apparent OH concentration in kinetic models.

In 2003, Bean *et al.* analyzed the integrated absorbances of the v_1 bands of the *cc*-HOONO and $HONO_2$ products with *ab initio* integrated cross-sections to determine a low-pressure *cc*-HOONO-to-$HONO_2$ product branching ratio of 7.5 K in 20 torr of buffer gas [4]. They also concluded that *trans-perp* HOONO consisted of a negligible fraction of total HOONO yield. This work was followed in rapid succession by *ab initio* molecular-orbital and transition-state theory calculations by Zhu *et al.*, which predicted the HOONO yield from reaction (2) to be less than 2.5% at 300–400 K and 1 atm, reaching 12.2% and 9.1% at 300 and 400 K respectively and 3.8×10^4 torr pressure [7].

Most recently, a multiwell, multichannel master-equation approach has been used by Golden *et al.* to estimate the fractional yield and decomposition rate of HOONO in the earth's atmosphere as a function of altitude (Figure 3), which predicted the upper-bound HOONO yields near the tropopause to be approximately 20% [8]. Experimental verification of the HOONO/$HONO_2$ branching ratio would greatly improve current models of the lower atmosphere [2].

An extension of the works of Bean *et al.* to detect and study *cc*-HOONO under atmospherically important conditions using cavity ringdown spectroscopy is currently in its earliest stages. An overview of the expected experimental apparatus is shown in Figure 4. In past experiments, a microwave generator has been used to form hydrogen atoms, which were then mixed with NO_2 to produce OH radicals by the reaction

$$H + NO_2 \longrightarrow OH + NO. \qquad (3)$$

However, this method is shown to be infeasible at high pressures [4]. Thus, the formation of OH through the laser-photolysis reactions

$$\begin{aligned} O_3 + h\nu &\longrightarrow O_2 + O(^1D) \qquad (4) \\ O(^1D) + H_2 &\longrightarrow OH + H \qquad (5) \end{aligned}$$

has been proposed.

The present work is a theoretical study that focuses on the reaction kinetics at various pressures to determine the necessary initial concentrations for HOONO detection under atmospheric conditions. Reaction kinetics reveal significant information about the concentration and activity of key components in atmospheric models. Also, the study of reaction kinetics can significantly simplify the number of reactions that must be considered in assessing the atmospheric fates of particular species of interest. The predicted values will be used to evaluate the feasibility of using laser photolysis to create a detectable concentration of HOONO.

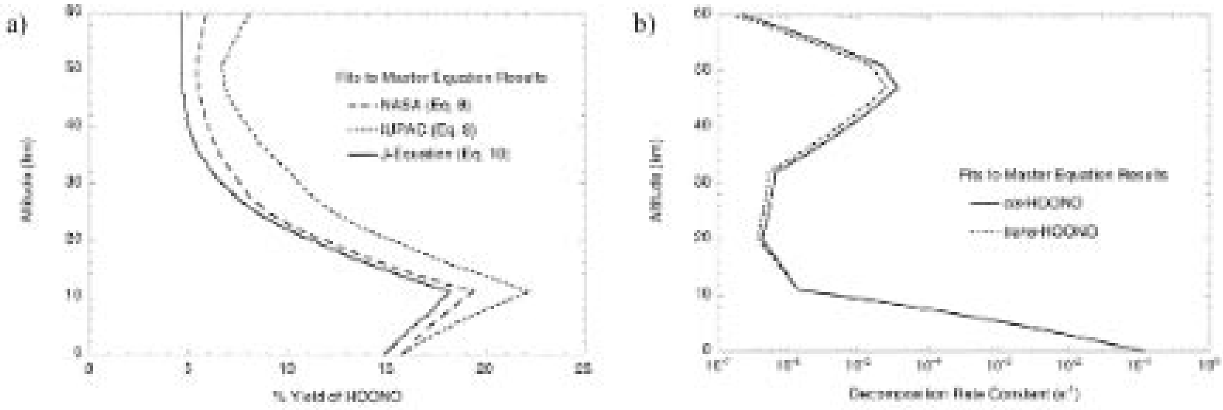


Figure 3: a) Predicted HOONO/$HONO_2$ branching as a function of altitude, based on the U.S. Standard Atmosphere, 1976 [8]. b) Predicted rate constants for thermal decomposition of *cc*- and *tp*-HOONO as a function of altitude, based on the U.S. Standard Atmosphere, 1976 [8].

2 Materials and Methods

A temperature- and pressure-dependent model of the photochemical reactions in *cc*-HOONO synthesis using laser photolysis was coded and analyzed using Facsimile, an integrating program which has been demonstrated to be a useful model tool for the development of models and analysis involving mixed systems of kinetic reactions. The model consisted of 33 chemical reactions.

The rate constants of the reactions are taken from NASA publication JPL02-25 [10], with the exception of reactions involving HOONO or HNO_3, whose rate coefficients have undergone significant revision since the release of JPL02-25. The rate constants for the exceptions can be found in the most recent JPL evaluation [11]. Furthermore, as no previous research has provided insight into the reaction $HOONO + OH \longrightarrow H_2O + NO_2 + O_2$, the rate constant is assumed to be similar to that of the reaction between HNO_3 and OH specifically for the sake of this model.

The chemical reactions were analyzed to trace the concentrations of various species over time under standard temperature and pressure to eliminate the reactions that are too slow to be of importance. Several species (HONO, HOONO, OH, NO_2, HNO_3, and $HONO_3$) were tagged and their concentrations were observed for the duration of the simulation to better understand their impact on product yields in the experimental cell. A heuristic approach was used to derive the necessary initial parameters to create a detectable amount of *cc*-HOONO. The model used in this experiment was constrained by several observations concerning the peak concentrations of H_2O, OH, HOONO, HNO_3, O_3, and HNO_2 [4]. Two additional models were constructed by the same method at 30 and 500 torr to provide insight into the photochemical processes over a wider atmospheric pressure range.

3 Results and Data

3.1 Sixteen Influential Reactions Identified

Figure 5 illustrates the average change in the concentrations of HONO, HOONO, OH, HNO_3, and $HONO_3$ when each reaction rate was doubled at 30, 100, and 500 torr. Negative concentration changes were accounted for by taking the percent change of the reverse reaction. The concentration changes ranged from -44.80% to 78.00%, with an average change of 3.15%. Sixteen reactions had an average percent change exceeding 1% and were identified as fast enough to be influential reactions. Eight additional reactions were noted to have significant effects at specific pressures.

The sixteen significant reactions[1] were further an-

---

[1] Reaction denotation in the Results section is independent of reaction denotation in the rest of the paper.

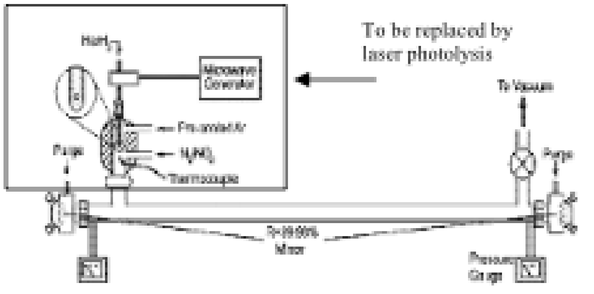


Figure 4: Overview of the expected experimental apparatus [4]. Laser photolysis will replace the microwave generator in the production of OH radicals. The OH radicals react with NO_2 to form HOONO and $HONO_2$, within 1 cm. Infrared absorption spectra of products are detected downstream in the ring-down cavity, formed by two high-reflectivity mirrors (dashed lines, $R = 99.98\%$ at 2.8 μm).

alyzed to determine their specific effects on each of the species; the results are shown in Table 1 at the end of the paper. The reaction $O(^1D)+N_2 \longrightarrow O+N_2$ was observed to have significant but opposing effects at the three different pressures. Two observations can be made from the data collected:

1. All quenching reactions (reactions 4–6) were significant.[2] Reaction 5, namely $O(^1D) + O_2 \longrightarrow O + O_2$, is significant at 30 and 100 torr.
2. Both channels of the reaction $OH + NO_2$ were significant, with the HOONO channel as the more influential reaction.

3.2 Effects of Changes in Initial Concentrations

Table 2 summarizes the effects of changing the initial species concentrations on the final HONO, HOONO, OH, HNO_3, and $HONO_3$ concentrations. Each initial concentration was doubled and its effects were monitored after 15 ms, when stability was reached in the model, to minimize error. The initial concentration of O'(D) was not taken into account due to its dependence on O_3 concentration in this model. Situations where the final concentration change was different at the three pressures were considered to be inconsequential for future experimentation due to a high probability of error.

3.3 Necessary Initial Concentration at 30, 100, and 500 Torr

Table 3 shows the initial concentrations of O'(D), O_2, O_3, H_2, NO_2, and NO necessary for the detection of *cc*-HOONO, as determined by Facsimile modeling. The final concentration was taken after 15 ms, when the concentrations of the species have reached a stable point. The expected *cc*-HOONO yield from $OH + NO_2$ is calculated to be 5.01%, 7.03%, and 12.3% at 30, 100, and 500 torr respectively.

3.3.1 Reaction Kinetics with Necessary Concentrations

Figure 6 illustrates the predicted species' concentration under the necessary initial concentrations as a function of time in 81 μs at 30 torr and 298 K. HOONO concentration reaches the detectable range (2×10^{13} mol/cm^3) at 53.5 μs and stabilizes at 2.08×10^{13} mol/cm^3 after 76.6 μs. The free radical OH reaches its peak concentration ($4.17x10^{14}$ mol/cm^3) at 0.7 μs and falls below HNO_3 concentration after 13.3 μs. The concentration of H_2O_2 after 81 μs is at a maximum of 9.23×10^{11} mol/cm^3. The relative concentrations of OH and H_2O_2 will become a strong point of discussion in this paper. The NO_2 product of OH + HOONO, labeled NO_2A, accounts for 0.01% of the total NO after 81 μs. Figure 7 shows the percent *cc*-HOONO yield from reaction 2 as a function of time in 81 μs at 30 torr and 298 K.

Figure 7 shows the species' concentration under the necessary initial concentrations and the percent *cc*-HOONO yield as a function of time in 81 μs at 100 torr and 298 K. HOONO concentration can be detected after 19.8 μs and stabilizes at 2.11×10^{13} mol/cm^3 after 46.0 μs. OH reaches peak concentration (2.78×10^{14} mol/cm^3) at 0.48 μs and falls below HNO_3 concentration after 4.95 μs. The final concentration of H_2O_2 is at a maximum of 6.26×10^{11} mol/cm^3. The NO_2 product HOONO is calculated to account for 0.003% of the total NO.

Figure 8 shows the results after 81 μs at 500 torr and 298 K. HOONO concentration is predicted to be detectable after 8.56 μs and stabilized at 2.08×10^{13} mol/cm^3 after 16.8 μs. The free radical OH will reach its peak concentration (1.44×10^{14} mol/cm^3) at 0.38

---

[2] Quenching reactions involve the collisional deactivation of an excited molecule. A common quenching molecule is $O(^1D)$, in which the electron is excited to the 4s orbital.

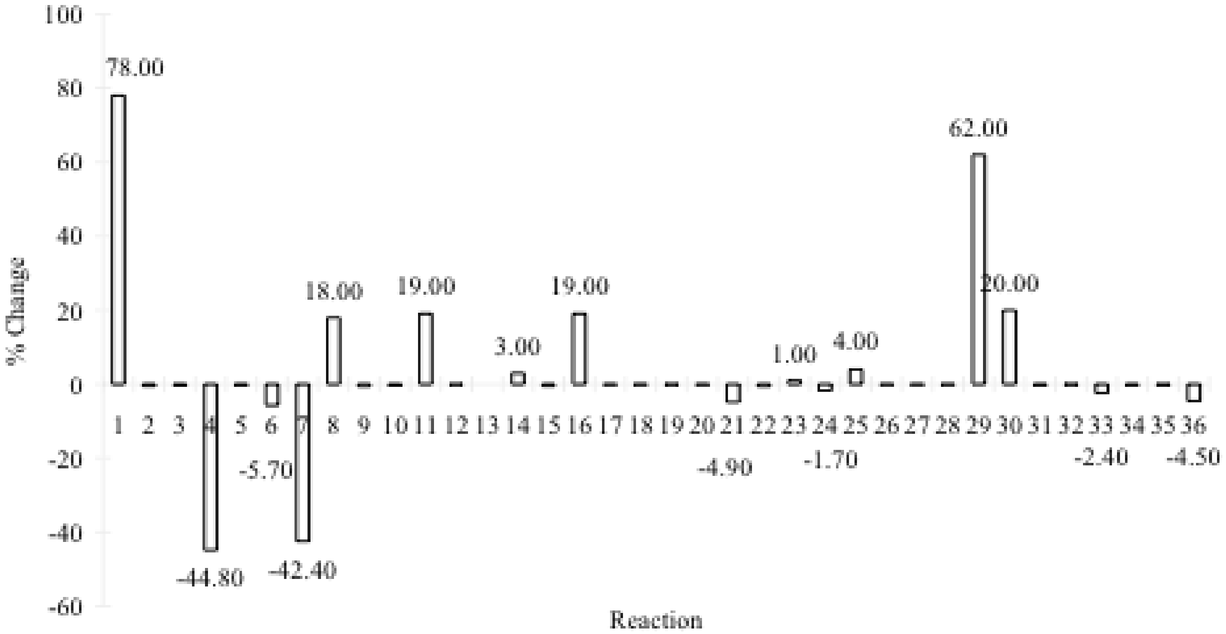


Figure 5: Average change in concentrations of HONO, HOONO, OH, HNO_3, and $HONO_3$ at 30, 100, and 500 torr when reaction rate is doubled. The data were taken from Facsimile modeling under the necessary concentration at 30, 100, and 500 torr and 298 K after 15 ms. Reactions that affected the concentrations by more than ±1% are considered significant.

μs and will fall below HNO_3 concentration after 2.42 μs. The concentration of H_2O_2 after 81 μs was at a maximum of 1.95×10^{11} mol/cm^3. NO_2A accounts for 6×10^{-4}% of the total NO after 81 μs.

3.3.2 Percent *cc*-HOONO Yield

Figure 9 compares the percent *cc*-HOONO yields as a function of time at 30, 100, and 500 torr and 278 K. The model appears to follow an exponential drop under all three pressures. The slope of the change in percent HOONO yields is observed to have a strong direct correlation with pressure. The difference between the initial and final percent *cc*-HOONO yields is calculated to be 0.88%, 0.32%, and 0.10% at 30, 100, and 500 torr respectively.

4 Discussion

The current model of NO_x reactions in the atmosphere consists of 36 reactions, 16 of which were identified as fast enough to be of importance for future atmospheric modeling at 30, 100, and 500 torr and 298 K. Both channels of the $OH + NO_2$ reaction were predicted to have a significant effect on the concentration of final products. This observation strongly supports the importance of HOONO in NO_x modeling. Our models also predict that collisional deactivation (quenching reactions) will significantly influence final concentrations. This phenomenon can be attributed to the relatively high pressures under which the simulation was conducted, which enhances the frequency of $O(^1D)$ collisions.

The percent cc-HOONO yields were determined at 30, 100, and 500 torr pressure to be 5.01%, 7.03%, and 12.3% respectively. Nizkorodovv and Wennberg reported a HOONO yield of 5 ± 3% at 253 K and 20 torr, which agrees with the Facsimile predictions when temperature and pressure adjustments are taken into account [6]. Burkholder *et al.* placed an upper limit of 5% on the HOONO yield; this limit was adjusted by Bean et al. to be 14% [4]. Golden *et al.* [8] used master-equation models to cap the HOONO yield at 20% near the tropopause ($\approx$ 100 torr). How-

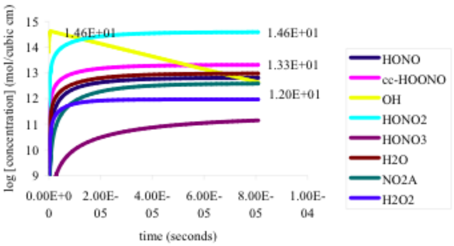


Figure 6: Predicted species concentration as a function of time, based on Facsimile modeling at 30 torr and 298 K.

ever, the high HOONO-yield limit is likely the result of the low temperature of the tropopause. Our model also agrees with their HOONO yield estimate of $\approx 15\%$ at 290 K and 95 torr.

Analysis of the reaction kinetics and species' concentrations under the predicted necessary initial conditions identified several potential issues that may occur in actual experimentation. The H_2O_2 concentration peaks were carefully studied to determine potential interference with the spectrum of *cc*-HOONO at the mid-infrared region. Six fundamental H_2O_2 bands were identified in the 3400–3800 cm^{-1} region, which are unlikely to interfere with detection of *cc*-HOONO (3306 cm^{-1}) [12]. The OH concentration may be high enough to interfere with detection at all three pressures by frequent interactions with neighboring species. The percent yield of NO_2 from HOONO was also closely traced to determine the relative amount of HOONO that is lost to the reaction $HOONO + OH \longrightarrow H_2O + O_2 + NO_2$. As less than a 1% loss was observed at all three pressures, this reaction can be considered insignificant in experimental determination of HOONO/HNO_3 branching ratios.

Effects of changes in the initial concentrations (Table 2) may be used to adjust initial concentrations for specific concerns in future experiments. Possible constraints may be found in the relationships of O3, O2, and O'(D) and the dependence of NO concentration on NO_2 concentration; however, these constraints were placed to facilitate NO and ozone collection in actual experimentation. Further attempts should be made to investigate the quantitative impacts of changing the initial concentrations to improve optimization efforts.

The necessary initial conditions derived from Facsimile modeling provide essential theoretical data for the use of laser photolysis to produce *cc*-HOONO in

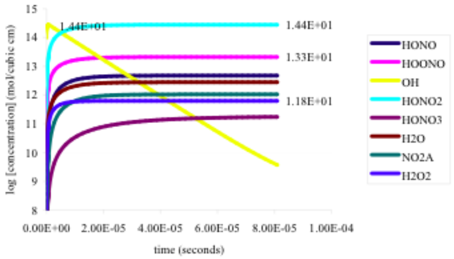


Figure 7: Predicted species concentration as a function of time, based on Facsimile modeling at 100 torr and 298 K.

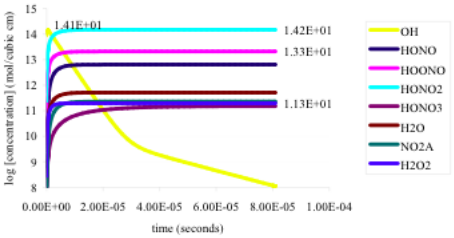


Figure 8: Predicted species concentration as a function of time, based on Facsimile modeling at 500 torr and 298 K.

spectroscopic detection at atmospherically relevant pressures. Predictions of HOONO yield show that it will increase significantly with pressure. An increase in this yield by a factor of two or three can greatly affect current atmospheric models. Thus experimental confirmation of HOONO yields would greatly contribute to our understanding of the lower atmosphere.

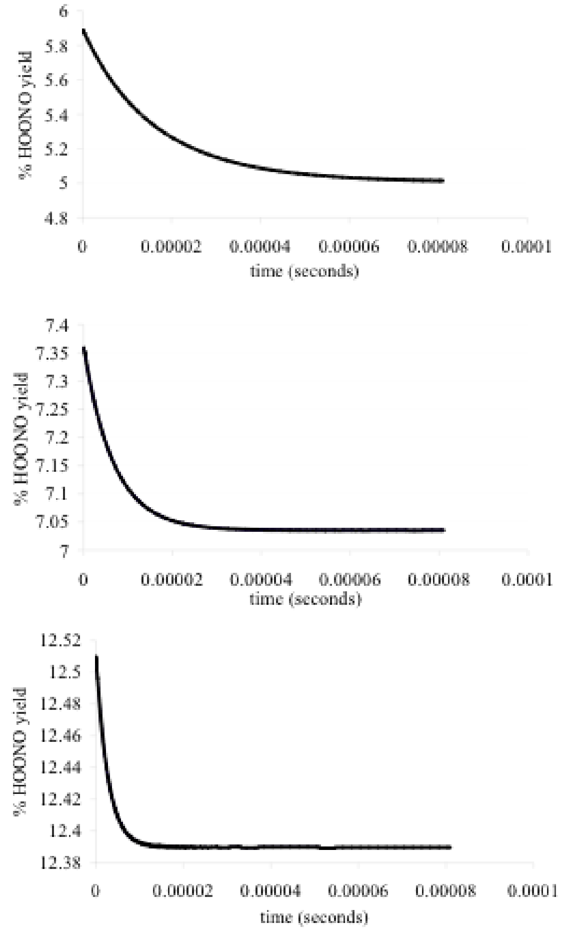


Figure 9: Predicted *cc*-HOONO yield from NO_2+OH reaction as a function of time, based on Facsimile modeling at (a) 30, (b) 50, and (c) 500 torr and 298 K.

5 Conclusion

This work has shown that it is feasible to utilize laser photolysis to produce a detectable concentration of *cc*-HOONO at atmospherically relevant pressures. Necessary initial concentrations for *cc*-HOONO detection were determined at 30, 100, and 500 torr. Further investigation identified reactions and initial species of importance to atmospheric modeling. The HOONO yields calculated from the Facsimile models are in very good agreement with previous experimental and theoretical studies. Analysis of reaction kinetics under calculated parameters provided insight into potential experimental concerns. Experimental confirmation from the detection and studies of HOONO under a wide range of atmospherically relevant pressures using cavity ringdown spectroscopy is expected to follow shortly.

6 Acknowledgments

I would like to extend my gratitude to Dr. Mitchio Okumura and Mr. Andrew Mollner of the California Institute of Technology for their guidance and support throughout this research experience. I would also like to thank Jacqueline Choi for her hard work and effort in helping me with this paper. Finally, I wish to express my appreciation for the Center for Excellence in Education, the California Institute of Technology, the staff of RSI, and all others who have contributed to my participation and enjoyment of the twenty-first annual Research Science Institute.

References

[1] D. McKee. *Tropospheric Ozone: Human Health and Agricultural Impacts.* Lewis Publishers, Boca Raton, FL (1993).

[2] B.J. Finalyson-Pitts and J.N. Pitts. *Chemistry of the Upper and Lower Atmosphere: Theory, Experiments, and Application.* Academic Press, San Diego, CA (1999).

[3] B.J. Drouin, J.L. Fry, and C.E. Miller. Rotational spectrum of *cis-cis* HOONO. Journal of Chemical Physics 120 (2004), 5505–5508.

[4] B.D. Bean, A.K. Mollner, S.A. Nizkorodov, G. Nair, M. Okumura, S.P. Sander, K.A. Peterson, and J.S. Francisco. Cavity ring-down spectroscopy of *cis-cis* HOONO and the HOONO/$HONO_2$ branching ratio in the reaction $OH + NO_2 + M$. Journal of Physical Chemistry A 107 (2003), 6974–6985.

[5] I.B. Pollack, I.M. Konen, E.X.J. Li, and M.I. Lester. Spectroscopic characterization of HOONO and its binding energy via infrared action spectroscopy. Journal of Chemical Physics 119 (2003), 9981–9984.

[6] S.A. Nizkorodov and P.O. Wennberg. First spectroscopic observation of gas-phase HOONO. Journal of Physical Chemistry A 106 (2002), 855–859.

[7] R.S. Zhu and M.C. Lin. *Ab Initio* study of the $HO_2 + NO$ reaction: prediction of the total rate constant and product branching ratios for the forward and reverse processes. Journal of Chemical Physics 119 (2003), 10667–10677.

[8] D.M. Golden, J.R. Barker, L.L. Lohr. Master equation models for the pressure- and temperature-dependent reactions $HO + NO_2 \longrightarrow HONO_2$ and $HO + NO_2 \longrightarrow$ HOONO. Journal of Physical Chemistry A 107 (2003), 11057–11071.

[9] B.J. Finalyson-Pitts and J.N. PItts. Tropospheric air pollution: ozone, airborne toxics, polycyclic aromatic hydrocarbons, and particles. Science 276 (1997), 1045–52.

[10] S.P. Sander, D.M. Golden, M.J. Kurylo, G.K. Monutgat, A.R. Ravishankara, C.E. Kolb, M.J. Molina, B.J. Finalyson-Pitts. *Chemical Kinetics and Photochemical Data for Use in Atmospheric Studies: Evaluation Number 14.* Jet Propulsion Laboratory (2003).

[11] S.P. Sander, D.M. Golden, M.J. Kurylo, G.K. Monutgat, A.R. Ravishankara, C.E. Kolb, M.J. Molina, B.J. Finalyson-Pitts. *Chemical Kinetics and Photochemical Data for Use in Atmospheric Studies: Evaluation Number 15.* Jet Propulsion Laboratory, unpublished.

[12] M. Pettersson, S. Tuominen, and M. Rasanen. IR spectroscopic study of H_2O_2, HDO_2, and D_2O_2 isolated in Ar, Kr, and Xe matrixes. Journal of Physical Chemistry A 101 (1996), 1166–1171.

Reactions	[HONO]	[*cc*-HOONO]	[OH]	[$\mathrm{HNO_3}$]	[$\mathrm{HONO_3}$]
1. $\mathrm{O(^1D) + H_2 \longrightarrow OH + H}$	↑	↑	↑	↑	↑
4. $\mathrm{O(^1D) + N_2 \longrightarrow O + N_2}$	No single trend between pressures				
6. $\mathrm{O(^1D) + NO_2 \longrightarrow O + NO_2}$	↓	↓	↓	↓	↓
7. $\mathrm{OH + NO_2 \longrightarrow HNO_3}$	↓	↓	↓	*No effect*	↓
8. $\mathrm{OH + NO \longrightarrow HONO}$	↑	*No effect*			
11. $\mathrm{OH + O_3 \longrightarrow HO_2 + O_2}$	*No effect*				↑
14. $\mathrm{O + NO_2 \longrightarrow O_2 + NO}$	↑	*No effect*			
16. $\mathrm{O + OH \longrightarrow O_2 + H}$	*No effect*				↑
21. $\mathrm{H + O_3 \longrightarrow OH + O_2}$	*No effect*				↓
23. $\mathrm{H + O_2 \longrightarrow HO_2}$	*No effect*				↑
24. $\mathrm{HO_2 + OH \longrightarrow H_2O + O_2}$	*No effect*				↓
25. $\mathrm{HO_2 + NO_2 \longrightarrow HONO_3}$	*No effect*				↑
29. $\mathrm{OH + NO_2 \longrightarrow HOONO}$	↓	↑	↑	↓	↓
30. $\mathrm{HOONO \longrightarrow OH + NO_2}$	*No effect*		↑	*No effect*	
33. $\mathrm{OH + HOONO \longrightarrow H_2O + O_2 + NO_2}$	*No effect*	↓	↓	*No effect*	
36. $\mathrm{NO + HO_2 \longrightarrow NO_2 + OH}$	*No effect*				↓

Table 1: Effects of influential reactions. The significant reactions are shown on the leftmost column, and the changes in the final concentration of the products after doubling the rate constant are represented by the arrows. Observations were made at 30, 100, and 500 torr and 298 K. Changes of magnitude within 0.1% were reported as no effect.

Increase in	HONO	*cc*-HOONO	OH	$\mathrm{HNO_3}$	$\mathrm{HONO_3}$	$\mathrm{H_2O}$	$\mathrm{NO_2A}$
$\mathrm{O_3}$	↑	↑	↑	↑	↑	↑	↑
$\mathrm{O_2}$	↓	↓	↓	↓	↑	*No trend observed*	
$\mathrm{H_2}$	↑	↑	↑	↑	↑	↑	↑
$\mathrm{NO_2}$	↓	↑	↓	↑	*No trend observed*	↓	↓
NO	↑	↓	↓	↓	↓	↓	↓

Table 2: Effects of doubling the concentration of the initial species. Data were calculated from Facsimile modeling after 15 ms at 30, 100, and 500 torr and 298 K. "No trend observed" indicates observations that differ with pressure and thus are not considered significant. Magnitude of changes was not taken into account.

Pressure (torr)	Initial Concentation (mol/cm^3)					
	O'(D)	$\mathrm{O_2}$	$\mathrm{O_3}$	$\mathrm{H_2}$	$\mathrm{NO_2}$	NO
30	1.70×10^{15}	1.70×10^{15}	1.70×10^{15}	4.00×10^{16}	4.00×10^{16}	2.13×10^{12}
100	1.80×10^{15}	1.80×10^{15}	1.80×10^{15}	7.50×10^{16}	4.00×10^{16}	2.25×10^{12}
500	5.40×10^{15}	5.40×10^{15}	5.40×10^{15}	6.40×10^{16}	4.00×10^{16}	6.75×10^{12}
Pressure (torr)	Final Concentration (mol/cm^3)					
	HONO	HOONO	OH	$\mathrm{HNO_3}$	$\mathrm{HONO_3}$	
30	6.52×10^{12}	2.08×10^{13}	1.23×10^{6}	3.94×10^{14}	2.05×10^{11}	
100	4.74×10^{12}	2.11×10^{13}	1.57×10^{6}	2.79×10^{14}	1.38×10^{11}	
500	6.37×10^{12}	2.08×10^{13}	2.67×10^{6}	1.47×10^{14}	1.52×10^{11}	

Table 3: Necessary initial concentration of species at 30, 100, and 500 torr. Initial concentrations were determined by Facsimile modeling. The final concentrations reported are the species' predicted concentrations after 15 ms.

Closure Properties of D_{2p} in Finite Groups

Po-Ling Loh

under the direction of
Dr. Michael Aschbacher
California Institute of Technology

Abstract

We discuss the generalized version of a question posed by the topologist E. Farjoun about closed embeddings of a finite group H in a finite group G. We study the case $H \cong D_{2p}$ for p an odd prime and determine a sufficient condition for $H = G$ when H is closed in G.

1 Introduction

Group theory is of importance in many fields of mathematics, including topology. The problem discussed in this paper is related to a problem posed by the topologist Farjoun.

Let H and G be objects in a category $\mathbf{C}$, and let $\iota \in \mathrm{Mor}(H, G)$. Farjoun studies morphisms ι such that for each $\varphi \in \mathrm{Mor}(H, G)$, there exists a unique $\varphi' \in \mathrm{Mor}(G, G)$ such that $\iota\varphi' = \varphi$. In other words, the following diagram commutes.

$$\begin{tikzcd} H \arrow[r, "\iota"] \arrow[dr, "\varphi"'] & G \arrow[d, "\varphi'"] \\ & G \end{tikzcd}$$

One special case arises when $\mathbf{C}$ is the category of groups and group homomorphisms, and ι is an inclusion map. We denote by $\mathrm{Hom}(H, G)$ the set of homomorphisms from H to G and write $\mathrm{End}(G) = \mathrm{Hom}(G, G)$. We say H is *closed* in G if $H \leq G$ and each $\varphi \in \mathrm{Hom}(H, G)$ extends uniquely to some $\varphi' \in \mathrm{End}(G)$.

Farjoun studies the properties of H that are preserved when H is closed in G. In particular, he asks the following question:

Question. *If H is finite, closed, and nilpotent, is $H = G$?*

It is known (cf. Remark 3.7) that if G is nilpotent and H is closed in G, then $H = G$. Therefore, Farjoun's question can be interpreted as asking whether nilpotence is preserved under the closure of finite groups. For more material related to Farjoun's question, we refer the reader to [3], [5], and [6].

In the discussion that follows, we will remove the condition that H be nilpotent and analyze general closed embeddings of finite groups in finite groups. In Section 2, we will consider the case when H and G are simple. Since H and G have only two normal subgroups apiece, we can more easily study the structure of elements of $\mathrm{Hom}(H, G)$ and $\mathrm{End}(G)$. Specifically, we will show that the closure of H in G is equivalent to three conditions.

If instead $H \cong D_{2p}$ for p an odd prime, then H has three normal subgroups: 1, H, and the derived subgroup $[H, H]$ of order p. Although the discussion is much more involved, we will arrive at the following result in Section 5:

Theorem. *If $H \cong D_{2p}$ and H is closed in a finite group G, then either $[G, G] = G^\infty$ or $H = G$.*

We also have the following corollary:

Corollary. *If $H \cong D_{2p}$ is closed in a finite solvable group G, then $H = G$.*

The condition in the corollary that G be solvable appears to be necessary, and we expect to be able to prove the following conjecture:

Conjecture. *For each $p > 3$, there exists a closed embedding of $H \cong D_{2p}$ in a finite group G, where G^∞ is a nonabelian finite simple group.*

In Sections 3 and 4, we will develop the lemmas necessary to prove the theorem. We will prove several facts about the relationship between H and G, first under the condition that (1) only the zero map 0_H

extends uniquely, and then under the stronger condition that (2) H is closed in G.

We assume that the reader is familiar with concepts in elementary group theory; see [1] and [4] for an introduction.

2 An Example

We begin with an example from Aschbacher's paper [2], supplying a proof of the assertion made there without proof.

Example. *Let H and G be simple. Then H is closed in G if and only if*

(i) $\mathrm{Aut}(G)$ *is transitive on subgroups of G which are isomorphic to H;*

(ii) $\mathrm{Aut}_{\mathrm{Aut}(G)}(H) = \mathrm{Aut}(H)$*; and*

(iii) $C_{\mathrm{Aut}(G)}(H) = 1$.

Proof. For each $\varphi \in \mathrm{Hom}(H,G)$, we have $\ker\varphi \trianglelefteq H$. Thus, the fact that H is simple implies $\ker\varphi = 1$ or H, and elements of $\mathrm{Hom}(H,G)$ are either injective homomorphisms or 0_H. Similarly, $\mathrm{End}(G) = \mathrm{Aut}(G) \cup \{0_G\}$.

Note that if $H = 1$, then H is closed in G if and only if $G = H = 1$, which is trivially equivalent to the given conditions. Thus, we need only consider the case when $1 < H \leq G$. Note also that if $G \neq 1$, the map 0_H is the only element of $\mathrm{Hom}(H,G)$ that extends to 0_G, and this extension is always unique.

It is easily verified that the given conditions are necessary for H to be closed in G. If (i) is not satisfied, there are some $K \leq G$ and some $\varphi \in \mathrm{Hom}(H,G)$ such that $H\varphi = K$ but φ does not extend to an element of $\mathrm{Aut}(G)$. If (ii) is not satisfied, there is some $\varphi \in \mathrm{Aut}(H) \leq \mathrm{Hom}(H,G)$ that is not the restriction of an element of $\mathrm{Aut}(G)$, so φ cannot extend to an element of $\mathrm{End}(G)$. If (iii) is not satisfied, there exists $1 \neq \alpha \in C_{\mathrm{Aut}(G)}(H)$, and both α and the identity map 1_G extend 1_H.

Next, we prove that the conditions are sufficient. Define $\theta : N_{\mathrm{Aut}(G)}(H) \to \mathrm{Aut}(H)$ by $\alpha\theta = \alpha|_H$ for each $\alpha \in N_{\mathrm{Aut}(G)}(H)$. Then $C_{\mathrm{Aut}(G)}(H) = \ker\theta = 1$, so θ is injective; hence, every element of $\mathrm{Aut}(H)$ in the image of θ extends uniquely to an element of $\mathrm{Aut}(G)$. Furthermore, (ii) implies θ is surjective. Thus, all elements of $\mathrm{Aut}(H)$ extend uniquely to elements of $\mathrm{End}(G)$.

To see that every element $\varphi \in \mathrm{Hom}(H,G)$ different from 0_H extends to at least one element of $\mathrm{End}(G)$, note that by (i), there exists $\psi \in \mathrm{Aut}(G)$ such that $H\varphi = H\psi$, since $H\varphi \cong H$. Then $\varphi\psi^{-1} \in \mathrm{Aut}(H)$ extends to $\rho \in \mathrm{Aut}(G)$, and $\rho\psi \in \mathrm{Aut}(G)$ induces φ on H.

To see that every element of $\mathrm{Hom}(H,G)$ different from 0_H extends to at most one element of $\mathrm{End}(G)$, suppose $\varphi', \varphi'' \in \mathrm{End}(G)$ satisfy $\varphi'|_H = \varphi''|_H$. Then $(\varphi'(\varphi'')^{-1})|_H = 1_H$, so $\varphi'(\varphi'')^{-1} \in C_{\mathrm{Aut}(G)}(H) = 1$ by (iii), implying $\varphi' = \varphi''$. Hence, the three given conditions are indeed equivalent to H being closed in G.

□

3 When 0_H Extends Uniquely

We now prove several facts under the more relaxed condition that the zero homomorphism of a subgroup extends uniquely.

Lemma 3.1. *0_H extends uniquely if and only if* $\mathrm{Hom}\big(G/\langle H^G\rangle, G\big) = 0$.

Proof. Suppose $\varphi \in \mathrm{End}(G)$ is an extension of 0_H. (Such a map exists because 0_G extends 0_H.) Since $\langle H^G\rangle$ is the normal closure of H in G, $H \leq \ker\varphi$ implies $\langle H^G\rangle \leq \ker\varphi$. Then φ factors through $G/\langle H^G\rangle$, and if $\pi : G \to G/\langle H^G\rangle$ is the canonical homomorphism, the map φ corresponds to a unique $\psi \in \mathrm{Hom}(G/\langle H^G\rangle, G)$ such that $\pi\psi = \varphi$. Hence, if 0_H extends uniquely to 0_G, we clearly have $\mathrm{Hom}(G/\langle H^G\rangle, G) = 0$, and if $\mathrm{Hom}(G/\langle H^G\rangle, G) = 0$, we have $\varphi = 0_G$. □

For the remainder of this section, assume $0_H \in \mathrm{Hom}(H,G)$ extends uniquely to $0_G \in \mathrm{End}(G)$.

Lemma 3.2. *If G^∞ is the last term in the derived series of G, then $G = \langle H^G\rangle G^\infty$.*

Proof. Recall that the derived series $\{G^i\}$ is defined by the rules that $G^0 = G$ and $G^{i+1} = [G^i, G^i]$ for $i \geq 0$. Furthermore G^∞ is the smallest normal subgroup of G such that G/G^∞ is solvable.

Now suppose $G \neq \langle H^G\rangle G^\infty$, and let the quotient map by G^∞ be denoted by a star. Then $\langle H^G\rangle \trianglelefteq G$ implies $\langle H^G\rangle^* \trianglelefteq G^*$, and since G^* is solvable, so is $G^*/\langle H^G\rangle^*$. Thus, if $X^*/\langle H^G\rangle^*$ is the last term in a composition series of $G^*/\langle H^G\rangle^*$, we have $|G^*/\langle H^G\rangle^* : X^*/\langle H^G\rangle^*| = |G : X| = p$, where p is a prime factor of $|G|$ and $\langle H^G\rangle \leq X \trianglelefteq G$. (The case $G = 1$ trivially implies $G = \langle H^G\rangle G^\infty$.)

By Cauchy's Theorem, there exists $Y \leq G$ with $|Y| = p$. Then G/X and Y are cyclic groups of the same order, so $G/X \cong Y$, and there exists an isomorphism $\varphi : G/X \to Y$. Since $\langle H^G\rangle \leq X$, there exists a homomorphism $\psi : G/\langle H^G\rangle \to G/X$. Then $0 \neq \psi\varphi \in \operatorname{Hom}\bigl(G/\langle H^G\rangle, G\bigr)$, which by Lemma 3.1 contradicts the fact that 0_H extends uniquely. $\square$

Remark 3.3. If G is solvable, then $G^\infty = 1$, so Lemma 3.2 states that if 0_H extends uniquely, then $G = \langle H^G\rangle$. The converse is also true in this case, by Lemma 3.1 and the fact that $\operatorname{Hom}\bigl(G/\langle H^G\rangle, G\bigr) = \operatorname{Hom}(G/G, G) = 0$.

Lemma 3.4. *If G is nilpotent and $P \leq G = \langle P^G\rangle$, then $G = P$.*

Proof. We prove for every $K \leq G$ that in fact $K \trianglelefteq\trianglelefteq G$. We induct on the index of K in G. If $|G : K| = 1$, then the claim is trivially true. If instead $K < G$, then the fact that G is nilpotent implies $K < N_G(K)$. Since $K \trianglelefteq N_G(K)$ and $|G : N_G(K)| < |G : K|$, we have by induction that $N_G(K) \trianglelefteq\trianglelefteq G$, so $K \trianglelefteq\trianglelefteq G$, as desired. From $P \trianglelefteq\trianglelefteq G$ and $G = \langle P^G\rangle$, we then obtain $G = P$. $\square$

Lemma 3.5. *Let $N \trianglelefteq G$, and let the quotient map by N be denoted with a star. Then $\langle H^G\rangle^* = \bigl\langle (H^*)^{G^*}\bigr\rangle$.*

Proof. Since $\langle H^G\rangle \trianglelefteq G$, we have $\langle H^G\rangle^* \trianglelefteq G^*$. Also, $H^* \leq \langle H^G\rangle^*$, so clearly $\bigl\langle (H^*)^{G^*}\bigr\rangle \leq \langle H^G\rangle^*$. Since $\langle H^G\rangle = \langle h^g : h \in H, g \in G\rangle$ and $(h^g)^* = (g^{-1}hg)^* = (g^{-1})^* h^* g^* = (h^*)^{g^*}$, we also have $\langle H^G\rangle^* \leq \bigl\langle (H^*)^{G^*}\bigr\rangle$, as desired. $\square$

Lemma 3.6. *If L^∞ is the last term in the lower central series of G, then $G = HL^\infty$.*

Proof. Recall that the lower central series $\{L^i\}$ is defined by the rules that $L^0 = G$ and $L^{i+1} = [L^i, G]$ for $i \geq 0$. Furthermore L^∞ is the smallest normal subgroup of G such that G/L^∞ is nilpotent.

Since $G^\infty \leq L^\infty$, Lemma 3.2 implies $G = \langle H^G\rangle G^\infty = \langle H^G\rangle L^\infty$. Let the quotient map by L^∞ be denoted by a star, and suppose $\bigl\langle (H^*)^{G^*}\bigr\rangle = K^* \trianglelefteq G^*$, where $L^\infty \leq K \neq G$. Lemma 3.5 shows that $\langle H^G\rangle^* = K^* \trianglelefteq G^*$, so $\langle H^G\rangle \leq K \trianglelefteq G$. Then $\langle H^G\rangle L^\infty \leq KL^\infty = K \neq G$, a contradiction. Hence $\bigl\langle (H^*)^{G^*}\bigr\rangle = G^*$.

Since G^* is nilpotent, Lemma 3.4 then implies $G^* = H^*$. Thus, for each $g \in G$, there exists $h \in H$ such that $gL^\infty = hL^\infty$, so $g \in hL^\infty$ and $G \leq HL^\infty$. Clearly $HL^\infty \leq G$, so $G = HL^\infty$, as desired. $\square$

Remark 3.7. If G is nilpotent, then $L^\infty = 1$, so Lemma 3.6 implies $H = G$.

Corollary 3.8. $G = H[G, G]$.

Proof. This follows from Lemma 3.6 and the fact that $L^\infty \leq [G, G]$. $\square$

4 When H is Closed in G

We now prove facts about closed embeddings of groups, which we later apply to the specific case $H \cong D_{2p}$.

Lemma 4.1. *Let $P \leq G$, and let 0_P and 1_P extend uniquely to elements of* $\operatorname{End}(G)$. *Then*

(1) $C_G(P) = Z(G)$, and

(2) if P is abelian, then $G = P$.

Proof. Clearly $Z(G) \leq C_G(P)$. For each $g \in C_G(P)$, we have $c_g|_P = 1_P$, where c_g is the inner automorphism on G induced by g. Since 1_P extends uniquely, we have $c_g = 1_G$, so $g \in Z(G)$ and $C_G(P) \leq Z(G)$, proving (1).

If P is abelian, then $P \leq C_G(P) = Z(G)$. Thus $G = C_G(P) = Z(G)$, so G is abelian and therefore nilpotent. Since 0_P extends uniquely, Remark 3.7 implies $G = P$, as desired. $\square$

For the remainder of this section, assume H is closed in G.

Lemma 4.2. *Given $\varphi \in$* $\operatorname{End}(G)$*, each element of* $\operatorname{Hom}(H\varphi, G\varphi)$ *extends to at most one element of* $\operatorname{End}(G\varphi)$.

Proof. Suppose there exist $\theta, \psi \in \operatorname{End}(G\varphi)$ such that $\theta|_{H\varphi} = \psi|_{H\varphi}$ and $\theta \neq \psi$. Then $\varphi\theta, \varphi\psi \in \operatorname{End}(G)$ and $(\varphi\theta)|_H = (\varphi\psi)|_H \in \operatorname{Hom}(H, G)$, but $\varphi\theta \neq \varphi\psi$, contradicting the fact that H is closed in G. $\square$

Lemma 4.3. *Let $\varphi \in$* $\operatorname{End}(G)$*, and let $H\varphi$ be abelian. Then $G\varphi = H\varphi$.*

Proof. By Lemma 4.2, each element of $\operatorname{Hom}(H\varphi, G\varphi)$ extends to at most one element of $\operatorname{End}(G\varphi)$. In particular, $0_{H\varphi}$ and $1_{H\varphi}$ extend to $0_{G\varphi}$ and $1_{G\varphi}$ respectively, so these extensions are unique. The result then follows from part (2) of Lemma 4.1. $\square$

Lemma 4.4. *Let Δ be a set of cyclic subgroups of G. Let Γ be the set of $K \trianglelefteq H$ such that H/K is isomorphic to an element of Δ. Then $H \cap [G, G] \leq \bigcap_{K\in\Gamma} K$.*

Proof. For each $K \in \Gamma$, there exist $X \in \Delta$ and $\varphi \in \operatorname{Hom}(H, G)$ such that $K = \ker\varphi$ and $X = H\varphi$. Then φ extends to $\varphi_K \in \operatorname{End}(G)$, and $K = H \cap \ker\varphi_K$. Also $H\varphi$ is cyclic and therefore abelian, so $G\varphi_K = H\varphi$ by Lemma 4.3.

Since $G\varphi_K$ is abelian, we have $[G, G] \leq \ker\varphi_K$. Thus $[G, G] \leq \bigcap_{K\in\Gamma} \ker\varphi_K$, and

$$H \cap [G, G] \leq \bigcap_{K\in\Gamma} H \cap \ker\varphi_K = \bigcap_{K\in\Gamma} K,$$

as desired. □

Lemma 4.5. *$H \cap [G,G] = [H,H]$.*

Proof. Clearly $[H,H] \leq H \cap [G,G]$. To prove the reverse containment, let the quotient map by $[H,H]$ be denoted with a star. Since H^* is abelian, we may write $H^* = R_1^* \times \cdots \times R_n^*$, where the R_i^*'s are cyclic subgroups of H^*. Let K_i be the preimage of $\prod_{j \neq i} R_j^*$, and let $r_i \in H$ be such that $R_i^* = \langle r_i^* \rangle$.

We first show that $|R_i^*|$ divides $|r_i|$. If $|R_i^*| = k$, then k is the smallest positive integer such that $r_i^k \in [H,H]$. If $|r_i| = l$, then $r_i^l = 1 \in [H,H]$, so $k \nmid l$ provides a contradiction to the minimality of k.

Now let $X_i = \left\langle r_i^{l/k} \right\rangle \leq G$. Apply Lemma 4.4 to $\Delta = \{X_1, \ldots, X_n\}$ and note that $H/K_i \cong R_i^* \cong X_i$, so $H \cap [G,G] \leq \bigcap_{K \in \Gamma} K \leq \bigcap K_i = [H,H]$, where the last equality follows from the fact that $\bigcap \prod_{j \neq i} R_j^* = 1$. Hence $H \cap [G,G] = [H,H]$, as desired. □

Corollary 4.6. *$G/[G,G] \cong H/[H,H]$.*

Proof. We use Lemma 4.5 and the fact that $G = H[G,G]$, from Corollary 3.8. Since $[G,G] \trianglelefteq G$, the Second Isomorphism Theorem implies $G/[G,G] = H[G,G]/[G,G] \cong H/H \cap [G,G] = H/[H,H]$. □

5 When H is Isomorphic to D_{2p}

Throughout Section 5, assume $H \cong D_{2p}$ for p an odd prime, and H is closed in a finite group G.

Lemma 5.1. *Let $L = [G,G]$. Then $|G/L| = 2$.*

Proof. Corollary 4.6 states that $G/L \cong H/[H,H]$. Since $[H,H]$ is the derived subgroup of order p in H, we have $|G/L| = 2$. □

Lemma 5.2. *If $\varphi \in \operatorname{Hom}(H,G)$ and $\psi \in \operatorname{End}(G)$ is the unique extension of φ, then exactly one of the following holds:*

(1) $\varphi = 0_H$ and $\psi = 0_G$;
(2) $\ker\varphi = [H,H]$ and $\ker\psi = L$; or
(3) φ is injective.

Proof. Since $\ker\varphi \trianglelefteq H$, the subgroup $\ker\varphi$ is one of H, $[H,H]$, and 1. The first case implies $\varphi = 0_H$, which extends uniquely to 0_G, giving (1).

In the second case, $H\psi = H\varphi \cong H/[H,H] \cong \mathbb{Z}/2$ is abelian, so Lemma 4.3 implies $G/\ker\psi \cong G\psi = H\psi$ is abelian as well. But G/L is the largest abelian quotient of G, so $L \leq \ker\psi$. Lemma 5.1 gives $|G/L| = 2$; since $\ker\psi \neq G$, we must have $\ker\psi = L$, giving (2).

The third case, $\ker\varphi = 1$, is clearly equivalent to (3). □

Lemma 5.3. *If $\psi \in \operatorname{End}(G)$ and $G\psi \cong H$, then $G = H$.*

Proof. Let $K = \ker\psi$. Then $G/K \cong G\psi \cong H$, so $|G| = |K||H|$. Since $|G/K| \neq 2$, necessarily $K \neq L$. Also $\psi \neq 0_G$, so Lemma 5.2 implies $\psi|_H$ is injective. Hence $K \cap H = \ker(\psi|_H) \cap H = 1$. Then $KH \leq G$ and $|KH| = |K||H| = |G|$ imply $G = KH$. Since $K \trianglelefteq G$, the group G is the semidirect product of K and H, and the map $\theta : G \to H$ defined by $\theta(kh) = h$ for any $k \in K$ is a homomorphism. But then $\theta|_H = 1_H$, so $\theta = 1_G$ and $kh = \theta(kh) = h$. Hence $K = 1$ and $G = H$. □

Lemma 5.4. *Let t be an involution in H. Then*

(1) $G = L\langle t \rangle$, and
(2) $L = L^\infty$.

Proof. Lemma 4.5 gives $H \cap L = [H,H]$, so $L \cap \langle t \rangle = 1$. Since $L \trianglelefteq G$, we have $L\langle t \rangle \leq G$. But $|G|/|L| = 2$ from Lemma 5.1, so $|L\langle t \rangle| = |L||\langle t \rangle| = |G|$ and $G = L\langle t \rangle$, giving (1).

From Lemma 3.6, we have $G = HL^\infty$. Furthermore $L^\infty \leq L$, so $H \cap L^\infty \leq H \cap L$. Since $|H \cap L| = p$, we have $|H \cap L^\infty| = 1$ or p. But $[H,H] = L^\infty(H) \leq L^\infty$ and $[H,H] \leq H$, so $|H \cap L^\infty| \neq 1$. From Corollary 3.8, we also have $G = HL$. Then $|H||L|/|H \cap L| = |G| = |H||L^\infty|/|H \cap L^\infty|$ implies $L^\infty = L$, giving (2). □

Lemma 5.5. *Let the quotient map by $[L,L]$ be denoted by a star. Then $L^* = [L^*, t^*]$. In particular, each element of L^* is inverted by t^*.*

Proof. Under the quotient map by $[L,L]$, the results of Lemma 5.4 become $G^* = L^*\langle t^* \rangle$ and $L^* = (L^\infty)^*$. Since $(L^\infty)^* = [(L^\infty)^*, G^*]$, we then have $L^* = [L^*, G^*] = [L^*, L^*\langle t^* \rangle]$. Furthermore L^* is abelian, so $L^* = [L^*, \langle t^* \rangle] = [L^*, t^*]$.

Now $t \notin [L,L]$ is an involution, so t^* is also an involution. If $(l^*)^{-1}t^*l^*t^*$ is a generator of L^*, then $t^*((l^*)^{-1}t^*l^*t^*)t^* = t^*(l^*)^{-1}t^*l^*$, so each generator of L^* is inverted by t^*. If $l_1^*, l_2^* \in L^*$ are both inverted by t^*, then $t^*(l_1^*l_2^*)t^* = (t^*l_1^*t^*)(t^*l_2^*t^*) = (l_1^*)^{-1}(l_2^*)^{-1} = (l_2^*)^{-1}(l_1^*)^{-1} = (l_1^*l_2^*)^{-1}$, so $l_1^*l_2^*$ is also inverted by t^*. Hence, all elements of L^* are inverted by t^*. □

Lemma 5.6. *All elements of L^* have odd order.*

Proof. Let $L_0^* \leq L^*$ be the subgroup of elements of odd order, and let S^* be a Sylow 2-group of L^*. Then

$L^* = L_0^* S^*$, so $G^* = L_0^* S^* \langle t^* \rangle$. Since $S^* \langle t^* \rangle$ is a 2-group (and hence nilpotent), we have $L_0^* \cap S^* \langle t^* \rangle = 1$. Also $L_0^* \trianglelefteq G^*$, so $G^*/L_0^* \cong S^* \langle t^* \rangle$. Then G^*/L_0^* is nilpotent and $L^* = L^{\infty *} \leq L_0^*$, so $L_0^* = L^*$, as desired. ☐

Lemma 5.7. *Either* $[G,G] = G^\infty$ *or* $H = G$.

Proof. Note that $L^* = 1$ implies $L = [L,L]$, which is equivalent to $[G,G] = G^\infty$. Lemma 5.5 shows that t^* inverts L^*, so t^* acts on each subgroup of L^*. Thus, if $L^* \neq 1$, there exists a maximal subgroup $M^* < L^*$ with $M^* \trianglelefteq G^*$. Let M be the preimage of M^* in G, and let the quotient map by M be denoted by a bar. By Lemma 5.6 and the fact that every maximal subgroup of an abelian group has prime index, we have $|L : M| = |L^* : M^*| = q$ for some odd prime q, and $|\bar{G}| = 2q$.

If we now pick $l \in L \backslash M$, we have $M \trianglelefteq L$ and $\langle l \rangle \leq L$, so $M \langle l \rangle \leq L$. Since M is a maximal subgroup of L, we also have $M \langle l \rangle = L$. Lemma 5.4 gives $G = L \langle t \rangle = \langle L, t \rangle$, which implies $\bar{G} = \langle \bar{l}, \bar{t} \rangle$. Since t^* is an involution and l^* is inverted by t^* (by Lemma 5.5), also $\bar{t}$ is an involution and $\bar{l}$ is inverted by $\bar{t}$. Thus $\bar{G} \cong D_{2q}$.

Now $|\bar{G}|$ is odd, so $\bar{t}^{\bar{l}}$ is an involution in $\bar{G}$. Hence $\bar{G} = \langle \bar{t}^{\bar{l}}, \bar{t} \rangle$ and $|\langle \bar{t}^{\bar{l}} \bar{t} \rangle| = q$. Then q divides $|t^l t|$, and there exists $y \in \langle t^l t \rangle$ such that $|y| = q$. Furthermore $t(t^l t)t = t(l^{-1} t l t)t = (t t^l)^{-1}$, so all elements of $\langle t^l t \rangle$, including y, are inverted by t. Hence $E = \langle t, y \rangle \cong D_{2q} \cong \bar{G}$, and there exists a surjective homomorphism $\psi : G \to E$, factoring through G/M, with $\ker \psi = M$.

We now apply Lemma 5.2 to $\psi \in \mathrm{End}(G)$ and $\varphi = \psi|_H$. Clearly $\psi \neq 0_G$ and $\ker \psi = M \neq L$, so φ is an injection. Then $H\varphi \leq E$ implies $2p | 2q$, so $p = q$ and $E \cong H$. Combining ψ with this isomorphism provides $\theta \in \mathrm{End}(G)$ with $G\theta \cong H$, so $G = H$ by Lemma 5.3, completing the proof. ☐

Corollary 5.8. *If* G *is solvable, then* $H = G$.

Proof. Since $D_{2p} \leq G$, the group G is not abelian. Hence $G^\infty = 1 \neq [G,G]$, and $H = G$ by Lemma 5.7. ☐

6 Conclusion

We have discussed general closed embeddings of finite groups and studied the specific cases when both H and G are simple and when $H \cong D_{2p}$. We have also proved several facts about embeddings of finite groups under the weaker condition that 0_H extends uniquely. In the future, we hope to apply similar methods to study closed embeddings of other classes of finite groups.

7 Acknowledgments

The author would like to thank Dr. Michael Aschbacher of the California Institute of Technology for his patience and dedication while supervising her project. She also extends a special thank-you to Dr. I. Martin Isaacs of the University of Wisconsin-Madison for introducing her to group theory. Many thanks to the Center for Excellence in Education and the staff of the Research Science Institute for making this research possible.

References

[1] M. Aschbacher. *Finite Group Theory.* 2nd ed. Cambridge University Press (2000).

[2] M. Aschbacher. On a question of Farjoun. To appear in Proceedings of the Gainsville Conference on Finite Groups.

[3] C. Casacuberta. On structures preserved by idempotent transformations of groups and homotopy types. In: *Crystallographic Groups and Their Generalizations (Kortrijk, 1999).* Contemporary Mathematics 262, American Mathematical Society, Providence, RI (2000), 39–68.

[4] D.S. Dummit and R.M. Foote. *Abstract Algebra.* 2nd ed. John Wiley and Sons, Inc., Hoboken, NJ (1999).

[5] A. Libman. Cardinality and nilpotency of localizations of groups and G-modules. Israel Journal of Mathematics 117 (2000), 221–237.

[6] J.L. Rodriguez, J. Scherer, and A. Viruel. Nonsimple localizations of finite simple groups. To appear in Journal of Algebra.

The Effects of Atmospheric Turbulence on Simulated Starlight in a Low-Pressure Telescope Environment

Renee Park

under the direction of
Dr. Benjamin Lane
Massachusetts Institute of Technology

Abstract

We investigated the feasibility of using nulling-interferometry techniques in the stratosphere to search for extrasolar planets. Studies have shown wavefront distortions from atmospheric turbulence to be too large to obtain clear ground-based images. Wavefront errors caused by atmospheric turbulence in the stratosphere are on the scale of 1 angstrom [4]; however, the level of the wavefront error inside a telescope tube is unknown. We used Kite, a system level testbed created by the Jet Propulsion Laboratory for the External Metrology System of the Space Interferometry Mission, to simulate atmospheric turbulence conditions inside a telescope tube in the stratosphere. We obtained 40 minutes of laser-metrology gauge readings in relatively stabilized temperatures at 5 mb. The data was analyzed using adaptive optics and blanking techniques to reduce the level of instrument noise. The integrated total error was approximately 1.849 angstroms, validating the possibility of applying nulling-interferometry techniques without sending a telescope into space.

1 Introduction

The first extrasolar giant planet, 51 Pegasi b, was discovered in 1995 using Doppler spectroscopy [7]. Since then, scientists have endeavored to directly detect the presence of planets outside our solar system. However, the small angular separation and high brightness contrast characteristic of star-planet systems has thus far prevented the use of direct detection of light from extrasolar planets. From 10 pc away, the planets from Mercury to Saturn are within 1” from the Sun. For these planets, the Sun-to-planet contrast ratio (the intensity of starlight versus the intensity of the light from a planet) exceeds 10^9. While the angular separation cannot be adjusted, it is possible to lower the contrast ratio by observing in infrared, where the contrast ratio is smaller (10^3 to 10^6), or by finding a way to cancel the brightness of the star's light relative to its surroundings.

We are primarily interested in the optical methods of detection to provide direct images of the planets. Two techniques have proven promising in selectively attenuating the direct starlight to provide views of planets: coronagraphy and nulling interferometry [10]. Coronagraphy suppresses starlight by physically blocking the light from the star from the telescope detector. The residual starlight forms an Airy pattern of diffraction (Figure 1), and the method is only useful beyond several Airy radii away from the star as the residual light exceeds the light from the planet (Figure 2). Therefore, coronagraphy will be effective only for the closest stars. Nulling interferometry uses destructive interference to cancel light directly from a star. The residual pattern created by a nulling interferometer falls off faster with angle than does the residual from a coronagraph. This allows one to detect planets close to the star, resulting in a significant increase in the number of observable targets and/or allowing one to use a significantly smaller telescope (factor of two).

The basis of nulling interferometry involves sampling an incoming wavefront with two apertures (telescopes), and adjusting internal paths such that when their wavefronts are combined, the two electric-field vectors are exactly out of phase, causing them to cancel (Figure 3).

Until now, the use of nulling interferometry has been considered feasible only outside Earth's atmosphere, which is an expensive undertaking. Atmo-

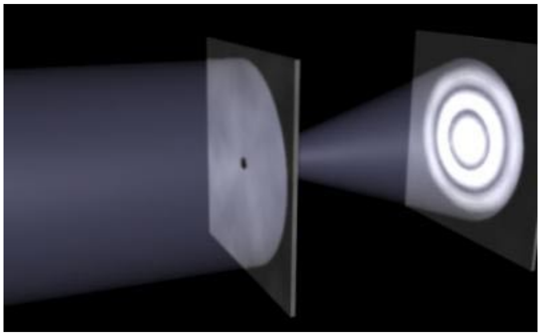


Figure 1: An Airy pattern forms when a light source passes through one hole. Nulling interferometry relies on two holes, forming a fringe pattern (a pattern of bars) [11].

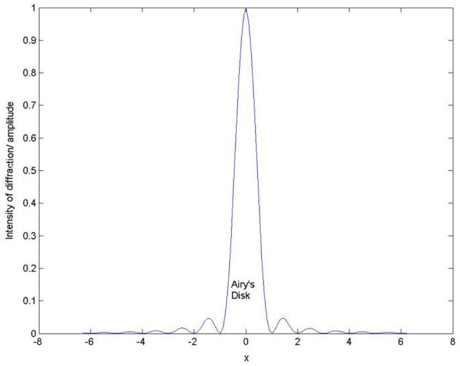


Figure 2: In an Airy pattern, the intensity is higher closer towards the light source (almost like a side-view). According to the scale of this diagram, if a planet is within 1 unit from its star, the planet will not be visible even after using coronagraphy [3].

spheric turbulence distorts incoming light and is a major contributor to the difficulty of optical and infrared interferometry from the ground. Many of the key design parameters of an interferometer are driven largely by the boundary conditions set by the atmosphere [10]. However, researchers have recently proposed that nulling interferometry could be used by telescopes transported on balloons. Balloons are capable of taking an instrument to the stratosphere, higher than about 99.5% of the atmosphere and a slightly larger part of the atmospheric turbulence. Past research shows that the wavefront error caused by the remaining atmospheric turbulence is on the scale of 10^{-10} m [4]. The major open question regarding the use of a balloon is what level of atmospheric turbulence is caused inside the telescope tube by the heat of the instrument at stratospheric altitudes.

To test the level of turbulence, we simulated conditions at the stratosphere using a vacuum chamber and measured the residual optical turbulence.

If it is proven that nulling interferometry can be applied successfully from balloons, then these results will serve to validate wavefront control and nulling technology necessary for Terrestrial Planet Finder missions currently being proposed by NASA. This method could also be used to survey stars not yet viewed by other techniques or avoided for other difficulties with accuracy.

2 Materials and Methods

2.1 Experimental Design

Data was collected using Kite, a testbed designed by the Jet Propulsion Laboratory for the External Metrology System of NASA's Space Interferometry Mission [1]. The original purpose of Kite was to test laser-metrology gauges in simulated space conditions. Laser-metrology gauges monitor internal and external optical-path distances required for astrometric measurements, using beam launchers to direct a metrology beam at a target and collect the returning laser light. The gauges are used for high-bandwidth sensing of phasing errors. We used metrology gauges inside a vacuum chamber that was adjusted to a pressure level of 5 millibars (3.5 torr), similar to that of the Earth's stratosphere.

Inside the chamber were six fiducials defining the baseline vectors and metrology gauges measuring the distances between the fiducials; the absolute lengths between the fiducials, on the order of 10 microns,

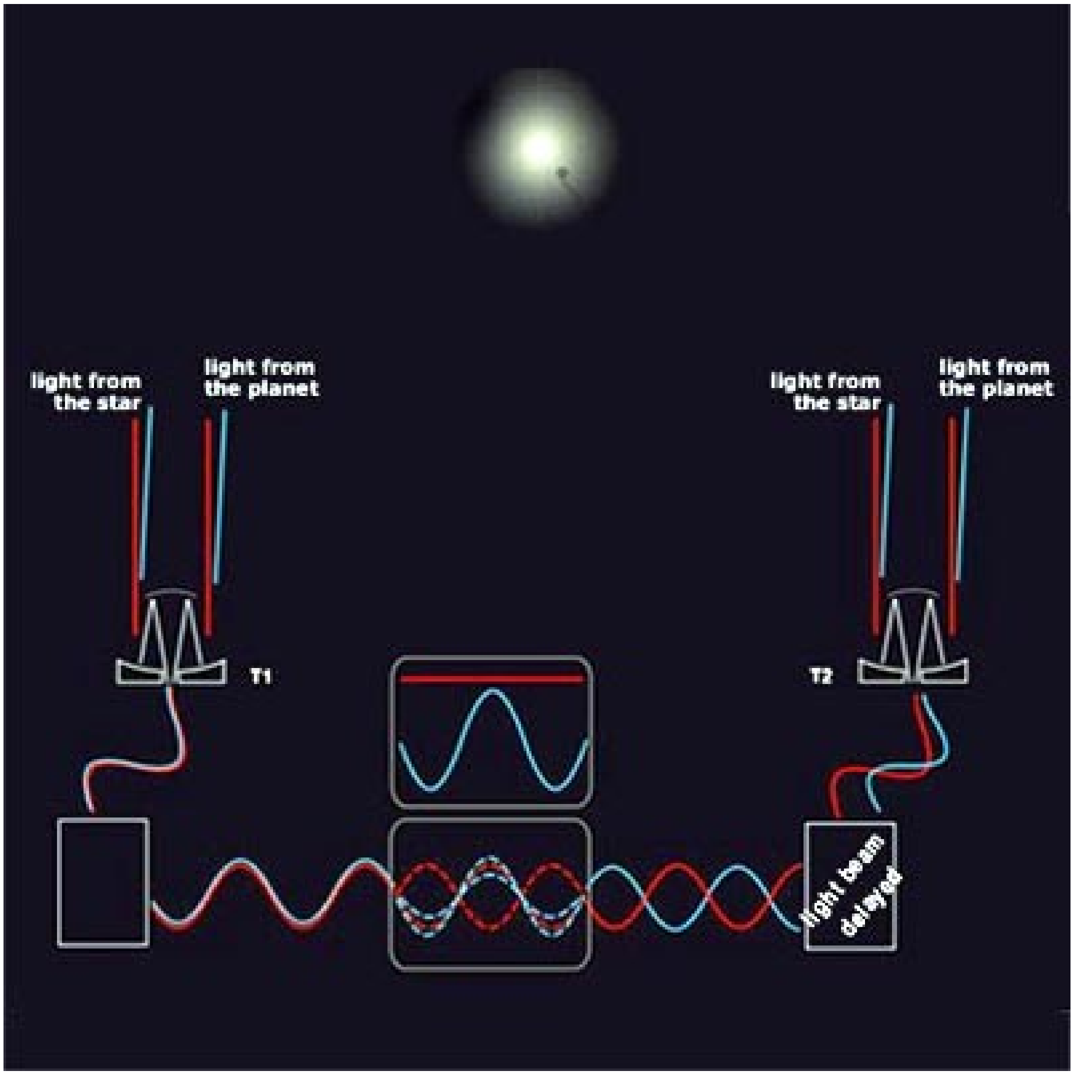


Figure 3: Nulling interferometry offsets two telescopes such that their images of a star cancel through destructive interference [2].

are known as 1.571615, 2.983117, 1.477939, 1.457499, 0.534182, 1.551161 for L1 to L6, respectively [1]. Four of the fiducials were double corner cubes and the other two were triple corner cubes all made of retroreflecting mirrors as shown in Figure 4; corner cubes are sets of mirrors arranged to reflect light from any direction directly back to its source. Kite measured the consistency between gauge measurements. Any path differences can be attributed to the remaining atmosphere or electrical noise in the metrology gauges. However, the electrical noise tends to occur at specific frequencies and can thus be filtered out.

2.2 Data Processing

2.2.1 Degree of Error

To evaluate the extent to which air turbulence affects traveling light waves, the predicted lengths measured by gauge 2 are calculated from the other five gauges and then compared to the actual readings from gauge 2. The points $((x_0, y_0), (x_1, y_1), (x_2, y_2), (x_3, y_3))$ in Figure 4 are placed on a coordinate plane where (x_0, y_0) is the origin and (x_2, y_2) is on the x-axis. The lengths are denoted l_i, where i runs from 1 to 6.

Consequently we have

$$\begin{aligned}(x_1, y_1) &= \left(\frac{l_5^2 - l_1^2 + l_3^2}{2l_5}, -\sqrt{l_3^2 - x_1^2}\right), && (1)\\ (x_2, y_2) &= (l_5, 0), && (2)\\ (x_3, y_3) &= \left(\frac{l_5^2 - l_4^2 + l_6^2}{2l_5}, -\sqrt{l_6^2 - x_3^2}\right). && (3)\\ & && (4)\end{aligned}$$

Thus, we find the predicted length between the top and bottom fiducials, whose actual length was read by gauge 2, can be expressed by the distance equation $l_2' = \sqrt{(x_1 - x_3)^2 + (y_1 - y_3)^2}$. The residual, or the difference between the predicted and measured length, is

$$\Delta \equiv l_2 - l_2'. \qquad (5)$$

Note, however, that for equations 1 through 4, the length is the absolute length, but the Kite instrument is only capable of measuring relative motions. Therefore, a control step called the "absolute metrology" mode determines the absolute length down to a resolution of tens of microns. Once the initial lengths have been determined, we add them to the relative gauge measurements before analyzing the data.

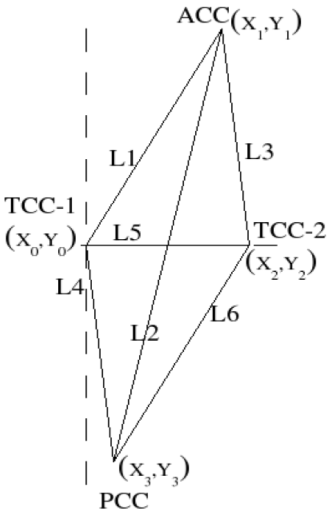


Figure 4: A schematic of the Kite instrument. Clockwise from left: triple corner cube 1 (TCC-1), articulating corner cube (ACC), triple corner cube 2 (TCC-2), and pointing corner cube (PCC). The "distances" between each pair of corners are labeled L1 to L6 [1].

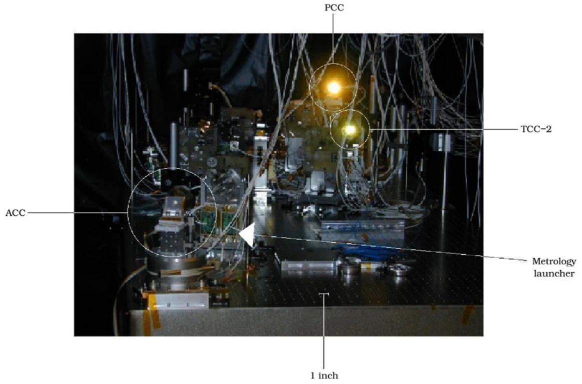


Figure 5: A photo of the Kite instrument. Note that the camera's flash is reflected directly back from the PCC and the TCC-2 because they are retroreflecting mirrors.[1]

2.2.2 Power Spectral Analysis

We evaluate the magnitude of the gauge residual produced by atmospheric turbulence using power spectral densities (PSDs). PSDs are based on the Fourier transform, which takes wave readings and gives information about the frequencies in those readings. The transformation restates periodic data as the sum of various sine and cosine waves. Based on the Euler formulae

$$\begin{aligned}\cos\rho &= \frac{e^{i\rho} + e^{-i\rho}}{2}, && (6)\\ \sin\rho &= \frac{e^{i\rho} - e^{-i\rho}}{2i}, && (7)\end{aligned}$$

we can use the Fourier integral in Octave, a freely distributed software based on Matlab that is designed primarily for numerical computations, to obtain PSDs of the data. This integral is

$$F(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(t)e^{2\pi i\omega t}dt, \qquad (8)$$

where $f(t)$ is a function of time and $F(\omega)$ is a function of wavelength.

2.2.3 Adaptive Optics

Once PSDs are obtained, the data must be corrected in a number of ways to reduce background noise. Because the Kite instrument uses adaptive optics, noise at low frequencies can typically be attributed to its

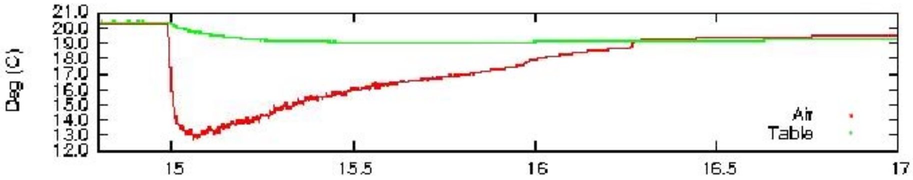


Figure 6: Temperature inside the vacuum chamber from 15 hours to 17 hours.

servo system. A servo system is an optical system that adapts to compensate for optical effects introduced by the medium between, in this case, a star and a detector. The system, usually operating at frequencies of 10 to 1000 Hz, actively tracks and corrects observational changes in a path. A simple servo system operates on one output parameter and is affected by some external disturbance. In order to control the system so that the result is as close as possible to the desired input, the sensor reads the output and corrects the command accordingly in a "feedback" system. However, although a servo system does correct most of the error caused by observational changes in the path, it also introduces some instrumental error that can be subtracted.

We want to quantify the effect of adding a notional servo loop to correct for the low-frequency path-length changes observed in the Kite data. We can use basic first-order servo theory to derive the servo noise rejection factor as a function of frequency, which is

$$\frac{1}{1+(f_c/f)^2}, \tag{9}$$

where f is the frequency and f_c is the constant closed-loop frequency, or how fast the servo system can correct the response time error.

2.2.4 Blanking Peaks

In addition to using adaptive optics to correct the wavefront error, a blanking technique can be used to remove error caused by instrument vibrations. Raw data will show a combination of sinusoidal waves when there is instrumental noise; the data will look periodic, which cannot be caused by atmospheric turbulence since atmospheric changes are random. On a PSD plot, the exact frequencies of the sinusoidal waves are seen as large peaks, which are then easily blanked by cutting off the peaks above a reasonable value and setting those frequencies to the average value of the entire plot. From previous experience, we expect instrumental noise to be at frequencies around 6 Hz and 10 Hz.

3 Results

We collected data using JPL's Kite instrument for over two hours. Since the temperature inside the vacuum did not stabilize while the pressure was being lowered, we only analyzed the final forty minutes of data (see Figure 6).

For clarity, we show the raw data detrended to one degree (Figure 7). The data has been decimated so that only one out of every hundred points are shown. No corner cubes are articulating and the vibrations only reflect instrument noise. However, note that gauges 1 through 3 have much stronger vibrations than gauges 4 through 6 because the first three gauges each see the articulating corner cube, which is excited by vacuum-tank modes causing about 50 nm vibrations.

The residual graph and the PSD of the residuals (Figures 8 and 9) both reflect a combination of sinusoidal functions affecting the data. As atmospheric turbulence is random, we know the periodic data is due to instrumental vibrations and can blank out their effects. By integrating the plots, illustrated by the cumulative sums on Figure 9, we know the total error is 1.264 nm and should be on the scale of a few angstroms after servo system corrections and blanking.

The servo system correction was applied with a calculated closed-loop frequency of 0.02 Hz. The result can be seen in Figure 10 where the PSD shows a rapid drop in the values in the lower frequencies. The cumulative sum at this point was approximately 3.314 angstroms.

Figure 10 also shows the plot after both servo corrections and blankings were applied. The large peaks around 6 Hz and 10 Hz in Figure 8 were blanked out, lowering all values above 10^{-8} to 10^{-10} for all

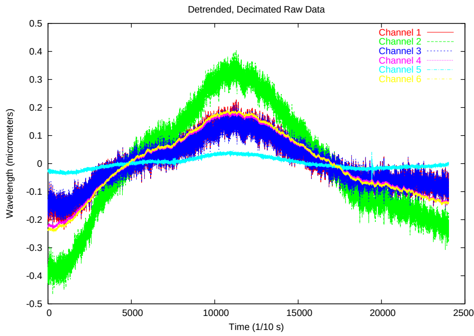


Figure 7: Detrended, raw, decimated data of all six channels.

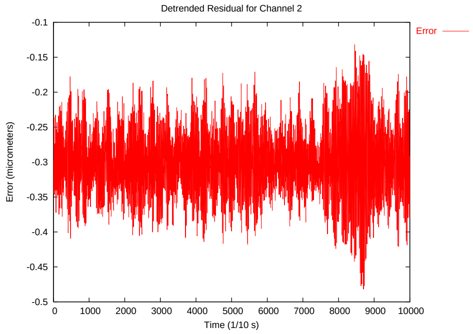


Figure 8: Detrended residual of first 10,000 points. Note the periodicities of the error.

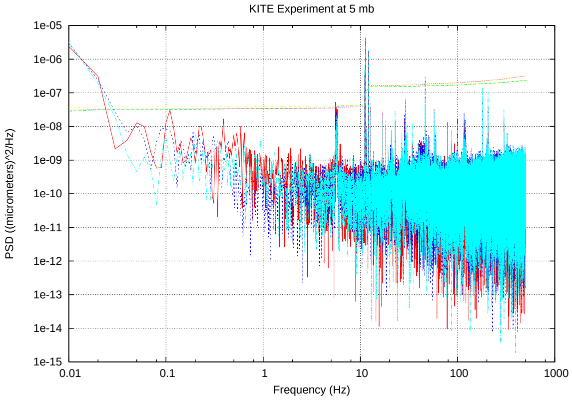


Figure 9: The PSD of three sets of 100,000 points from the residual and their cumulative sums.

frequencies greater than 1 Hz. The final cumulative error was 1.849 angstroms, which is a little less than twice the wavefront error found in the stratosphere [4]. Even beyond 10 Hz, the PSD shows large fluctuations. However, these points are random noise at low levels and do not amount to much integrated power.

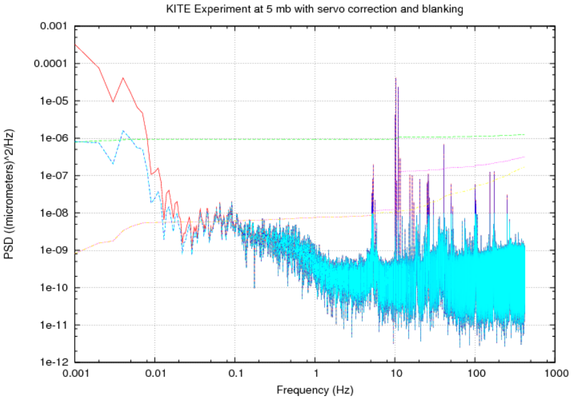


Figure 10: Three plots showing the original PSD (red), the PSD after the servo correction has been applied (dark blue), and the PSD after the servo correction and the blanking have been applied (light blue). Each plot is accompanied by their cumulative sums in green, pink, and yellow, respectively.

4 Conclusion

Using adaptive optics and blanking techniques, we have shown that in an environment approximating the conditions found in the stratosphere, the root-mean-square path-length error caused by turbulence over a path of 2.983 meters is 1.849 angstroms. We note also that the temperature difference between the air and the table was ±0.1 degrees Celsius. The level of turbulence typically scales with Δt (temperature) to the power of 1.2. Blanking techniques, instead of filtering techniques, were used because although results would be more accurate with filtration, the difference in the respective results would be small enough to be insignificant.

Our results are very promising as the error from atmospheric turbulence is only one-fifth of a wavelength. When spread out over an image of 1000 pixels, the error per pixel is only 0.0002. This error is sufficiently small that it will not mask the intensity of a planet's light.

We hope to use this data to support the use of balloon-borne telescopes for gathering data on extrasolar planets. This venture could be completed in 3 to 4 years, whereas building another Hubble-like telescope will not be completed for at least another decade.

5 Acknowledgments

I would like to thank my mentor, Dr. Benjamin F. Lane of the Massachusetts Institute of Technology, for all his invaluable help and guidance in this project, the researchers at Caltech who ran our experiment on the Kite instrument, and his colleagues Professor Edmund Bertschinger and Dr. Miguel Morales for their instruction on background information. I would also like to thank Balint Veto, my tutor at the Research Science Institute, for his support and comments. Finally, I would like to thank the Center for Excellence in Education for providing me the chance to work on this project.

References

[1] F.G. Dekens et al. Kite: status of the external metrology testbed for SIM. Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA (2004).

[2] European Space Agency. ESA to test the smartest technique for detecting extrasolar planets from the ground. Available at `http://www.esa.int/export/esaCP/ESAEQFF18ZC_Life_1.html` (2004/07/19).

[3] Glossary of PRIMA. Airy's disk. Available at `http://www.eso.org/projects/vlti/instru/prima/glossary_prima.html` (2004/07/19).

[4] D.P. Greenwood. Bandwidth specifications for adaptive optics systems. Journal of the Optical Society of America 67 (1977), 390–392.

[5] B. Lane. *High Precision Infra-red Stellar Interferometry*. Ph.D. thesis, 2003.

[6] B. Lane. *Prospects of a Balloon-borne Planet Imager*. Proposal to NASA (2004/04/19).

[7] M. Mayor and D. Queloz. A Jupiter-mass companion to a solar-type star. Nature 378 (1995), 355.

[8] L. Pallas. Harmonic analysis. Available at `http://math.ut.ee/~toomas_l/harmonic_analysis/Fourier/F.html` (2004/06/30).

[9] E. Serabyn. *Nulling Interferometry and Planet Detection*. Course notes, Michelson Summer School (1999).

[10] A. Quirrenbach. *Observing Through the Turbulent Atmosphere*. Course notes, Michelson Summer School (1999).

[11] The XMM Satellite Schoolpage. Why is diffraction so important? Available at `http://www.sr.bham.ac.uk/xmm/diffpage2.html` (2004/07/13).

On the Solvability of $\mathfrak{p}$-adic Diagonal Equations

Antoni Rangachev

under the direction of
Mr. Mohsen Bahramgiri
Massachusetts Institute of Technology

Abstract

This paper deals with some fundamental questions in the study of the diagonal diophantine equation $a_1x_1^k + \cdots + a_sx_s^k = 0$ over a finite extension K of the field $\mathbb{Q}_p$ of p-adic numbers, namely some new upper bounds on the number of variables needed to ensure its solvability.

1 Introduction

A basic problem in the study of diophantine equations is that of determining sufficient conditions for ensuring their solvability. Let p be a prime, let $\mathbb{Q}_p$ denote the field of p-adic numbers, and let K be an extension of $\mathbb{Q}_p$ of finite degree $n = [K : \mathbb{Q}_p]$. One of the fundamental questions in the theory of diophantine equations is, when does a diagonal equation, i.e., an equation of the form

$$a_1x_1^k + \cdots + a_sx_s^k = 0, \tag{1}$$

where the coefficients a_i are in the ring $\mathfrak{O}_K$ of integers of K, have a nontrivial solution over K? (By "nontrivial solution" we mean a nonzero vector $\mathbf{x} = (x_1, \dots, x_s) \in K^s$ satisfying (1).) When $K = \mathbb{Q}_p$, it is well known [5] that it suffices to have $s \geqslant k^2 + 1$. Generally, suppose $k = p^t m$, with $(m, p) = 1$. Let f be the residue class degree of K, and $d = (m, p^f - 1)$. The best known result in the case of arbitrary fields was established by Birch [2], who showed that for any K, it suffices to have $s \geqslant (2t+3)^k(d^2k)^{k-1}$. In 1996, Skinner [1] proved that when $k = p^t$ it suffices to have $s \geqslant k\big((k+1)^{\max(2t,1)} - 1\big) + 1$. The original formulation of Skinner's result is that the inequality above holds for every k. Unfortunately, it later turned out that there was an error in Skinner's proof. In an attempt to reconstruct this result we have obtained the following three results. In the first of these e denotes the ramification index of p in K, which is in any case at most n.

Theorem 1. *If $s \geq k((em+1)^{2t+1} - 1) + 1$, then any equation of the form* (1) *has a nontrivial solution over K.*

This general-purpose bound is qualitatively similar to that in Skinner's claim. The result is sharp when $K = \mathbb{Q}_p$ and $k = p - 1$.

If one is interested in fields K of small degree and increasingly large exponents k that are not highly divisible by p, the following bound is better than that of Theorem 1.

Theorem 2. *If $s \geq k(p^{n(2t+1)} - 1) + 1$, then any equation of the form* (1) *has a nontrivial solution over K.*

The final result shows that if the exponent k is fixed and indivisible by p, then the solubility threshold for diagonal equations is bounded independently of the field K.

Theorem 3. *If $(k, p) = 1$ and $s \geq k^2 + 1$, then any equation of the form* (1) *has a nontrivial solution.*

It would be interesting to prove a similar boundedness property without the restriction $(k, p) = 1$.

2 Notation and preliminaries

In order to describe our new results we need some notation, which we adopt from [1]. We denote by $\mathfrak{O}$ the ring of integers of K and let $\mathfrak{p} = (\pi)$ be its maximal ideal. Write f for the residue class degree of K and e for the ramification index of p in K, and let t and m be the integers such that $k = p^t m$ and $(m, p) = 1$. Let L be the maximal unramified subfield of K and $\mathfrak{o}$ its ring of integers. We recall that $\{1, \pi, \dots, \pi^{e-1}\}$ is an $\mathfrak{o}$-basis of $\mathfrak{O}$.

Let $\Gamma(k)$ be the least positive integer such that if $s \geq \Gamma(k)$, then any equation of the form (1) is solvable nontrivially over K. By $\Gamma_1(k)$ we denote the least positive integer such that any equation of the form (1) with $a_i \not\equiv 0 \pmod{\pi}$ for all i has a solution.

We say that $\mathbf{x}$ is a *nontrivial solution modulo* π^n if $\mathbf{x} = (x_1, \ldots, x_s) \in \mathfrak{O}^s$ is a solution of (1) modulo π^n and if additionally $x_j \not\equiv 0 \pmod{\pi}$ for some j. By $\Phi(k,n)$ and we denote the minimum s such that any equation (1) has a nontrivial solution modulo π^n. We shall write N for whichever of the integers $em+1$ and p^n enters into the bound we are proving.

2.1 The reduction lemma

The following lemma reduces the proof of our three main results to showing that $\Phi(k,e) \leq N$; recall that e is the ramification index of K over $\mathbb{Q}_p$.

Lemma 2.1.1 (adapted from Skinner [1]). *The functions Γ, Γ_1, and Φ satisfy the following inequalities:*

1. $\Gamma(k) \leq k\big(\Gamma_1(k) - 1\big) + 1$;
2. $\Gamma_1(k) \leq \Phi(k, 2et+1)$;
3. $\Phi\big(k,(v+1)e\big) \leq \Phi(k,e)\Phi(k,ve) \leq \Phi(k,e)^{v+1}$ *for every positive integer v;*
4. *if $\Phi(k,e) \leq N$, then $\Gamma(k) \leq k\big(N^{2t+1} - 1\big) + 1$.*

Proof. 1. Using the fact that every element of K can be written in the form $x = u\pi^{v_p(x)}$, where u is a unit, we can write all the coefficients a_i in the form $a_i = \pi^{r_i k + c_i} b_i$ for integers $r_i \geq 0$ and $0 \leq c_i < k$ and a $b_i \in \mathfrak{O}$ indivisible by π. If $s > k(c-1)$, then by the Pigeonhole Principle at least c of the c_i are the same. We may assume the corresponding indices to be $i = 1, \ldots, c$. Thus it suffices to find a nontrivial solution to the equation

$$b_1 x_1^k + \cdots + b_c x_c^k = 0, \quad \text{where } (b_i, \pi) = 1, \tag{2}$$

whose existence is guaranteed as long as $c \geq \Gamma_1(k)$.

2. Now we know that $a_i \not\equiv 0 \pmod{\pi}$ for all i. Put $r = 2et+1$. If $s \geq \Phi(k,r)$, then by the definition of $\Phi(k,r)$, there exists a nontrivial solution of (1) modulo π^r. Let $\mathbf{x} = (x_1, \ldots, x_s)$ be such a solution. We may assume that $x_1 \not\equiv 0 \pmod{\pi}$. Choose $y_2, \ldots, y_s \in \mathfrak{o}$ such that $y_i \equiv x_i \pmod{\pi^r}$, and let $d = \sum_{i=2}^{s} a_i y_i^k$. Since $a_1 x_1^k + d \equiv 0 \pmod{\pi^r}$, it follows from Hensel's Lemma that we can find $y_1 \in \mathfrak{o}$ such that $y_1 \equiv x_1 \pmod{\pi^r}$ and $a_1 y_1^k + d = 0$. Thus $\mathbf{y} = (y_1, \ldots, y_s)$ is a nontrivial solution of (1).

3. Let $h = \Phi(k, ve)$ and $l = \Phi(k,e)$ and for $j = 0, \ldots, l-1$ write

$$F_j(\mathbf{x_j}) = a_{jh+1} x_{jh+1}^k + \cdots + a_{(j+1)h} x_{(j+1)h}^k,$$

where $\mathbf{x}_j = (x_{jh+1}, \ldots, x_{j+1}h)$. Then (1) becomes

$$F_0(\mathbf{x}_0) + F_1(\mathbf{x}_1) + \cdots + F_{l-1}(\mathbf{x}_{l-1}) + \sum_{i=lh+1}^{s} a_i x_i^k = 0.$$

Thus, it suffices to find a nontrivial solution to the congruence

$$F_0(\mathbf{x}_0) + \cdots + F_{l-1}(\mathbf{x}_{l-1}) \equiv 0 \pmod{\pi^{(v+1)e}}. \tag{3}$$

By definition of $\Phi(k, ve)$, there exist nontrivial solutions $\mathbf{y}_j$ of the l equations

$$F_j(\mathbf{x}_j) \equiv 0 \pmod{\pi^{ve}}, \quad j = 0, \ldots, l-1.$$

Let $f_i = F_j(\mathbf{y}_j)$. Substituting $\mathbf{x}_j = t_j \mathbf{y}_j$ in (3), we get the new equation

$$f_0 t_0^k + \cdots + f_{l-1} t_{l-1}^k \equiv 0 \pmod{\pi^{(v+1)e}}, \tag{4}$$

where $f_j \equiv 0 \pmod{\pi^{ve}}$ for $0 \leq j < l$. From the definition of $\Phi(k,e) = l$, equation (4) has a nontrivial solution $\mathbf{t} = (t_0, \ldots, t_{\Phi(k,e)-1})$. Thus $\mathbf{y} = (t_0\mathbf{y}_0, \ldots, t_{\Phi(k,e)-1}\mathbf{y}_{\Phi(k,e)-1}, 0 \ldots, 0) \in \mathfrak{o}^s$ is a nontrivial solution of (1) modulo $\pi^{(v+1)e}$.

4. Because $\Phi(k,r)$ is nondecreasing in r, we deduce from statements 2 and 3 that $\Gamma_1(k) \leq \Phi(k, 2et+1) \leq \Phi(k,(2t+1)e) \leq \Phi(k,e)^{2t+1} \leq N^{2t+1}$, whence the desired inequality follows from statement 1. ☐

2.2 Chevalley-Warning Theorem

In this section we discuss some classical results concerning the solvability of equations over finite fields.

Let q be a power of a prime number p, and let $\mathbb{F}_q$ be a field with q elements. Consider the ring $\mathbb{F}_q[x_1, \ldots, x_n]$ of polynomials in n variables over $\mathbb{F}_q$.

In 1935 Artin conjectured that if $f(x) \in \mathbb{F}_q[x_1, \ldots, x_n]$ satisfies $f(0, \ldots, 0) = 0$ and $n > \deg f$, then f has at least one nontrivial zero in $\mathbb{F}_q^n$. Here we prove a generalization due to Chevalley and Warning.

Theorem 2.2.1 (*Chevalley-Warning*, cf., e.g., [9]). *Let $f_i \in K[x_1, \ldots, x_n]$ be polynomials in n variables such that $\sum \deg f_i < n$, and let V be the set of their common zeros in $\mathbb{F}_q^n$. Then*

$$\mathrm{Card}(V) \equiv 0 \pmod{p}.$$

Proof. We use the following lemma.

Lemma 2.2.1. *Let u be a nonnegative integer. The sum $S(X^u) = \sum_{x \in \mathbb{F}_q} x^u$ is equal to -1 if $u \geq 1$ and u is divisible by $q-1$; it is equal to 0 otherwise.*

Proof. If $u = 0$, all the terms of the sum are equal to 1; then $S(X^u) = q \cdot 1 = 0$ because $\mathbb{F}_q$ is of characteristic p. If $u \geq 1$ and u is divisible by $q-1$, we have $0^u = 0$ and $x^u = 1$ if $x \neq 0$. Hence $S(X^u) = (q-1) \cdot 1 = -1$.

Finally, if $u \geq 1$ and u is not divisible by $q-1$, the fact that the multiplicative group $\mathbb{F}_q^*$ of nonzero elements of $\mathbb{F}_q$ is cyclic of order $q-1$ shows that there exists $y \in \mathbb{F}_q^*$ such that $y^u \neq 1$. Then

$$S(X^u) = \sum_{x \in \mathbb{F}_q^*} x^u = \sum_{x \in \mathbb{F}_q^*} y^u x^u = y^u S(X^u),$$

whence $(1 - y^u)S(X^u) = 0$, or $S(X^u) = 0$. $\square$

Now put $P = \prod_\alpha (1 - f_\alpha^{q-1})$ and let $x \in \mathbb{F}_q^n$. If $x \in V$, all the $f_\alpha(x)$ are zero and $P(x) = 1$; if $x \notin V$, one of the $f_\alpha(x)$ is nonzero and $f_\alpha^{q-1} = 1$, hence $P(x) = 0$. If, for every polynomial f, we put $S(f) = \sum_{x \in \mathbb{F}_q} f(x)$, we have

$$\mathrm{Card}(V) \equiv S(P) \pmod{p}$$

and we reduce the claim to showing that $S(P) = 0$. Now the assumption $\deg f_\alpha < n$ implies that $\deg P < n(q-1)$; thus P is a linear combination of monomials $X^u = X_1^{u_1} \cdots X_n^{u_n}$ with $\sum u_i < n(q-1)$. It suffices to prove that $S(X^u) = 0$ for such monomials X^u, and this follows from the lemma since at least one u_i is smaller than $q-1$. $\square$

Corollary 2.2.1. *If $\sum \deg f_\alpha < N$ and the f_α have no constant term, then the f_α have a common nontrivial zero.*

2.3 Algebraic-number-theory lemmas

Before we are able to prove our results, we need some standard algebraic-number-theory properties of the ring $\mathfrak{O}$ and its maximal ideal $\mathfrak{p}$.

Lemma 2.3.1. *There exists an isomorphism*

$$\mathfrak{p}^n/\mathfrak{p}^{n+1} \cong \mathfrak{O}/\mathfrak{p}.$$

Proof. Let $\mathfrak{a} \neq 0$ be an arbitrary ideal of $\mathfrak{O}$ and $x \neq 0$ an element in $\mathfrak{a}$ with the smallest possible value $v(x) = n$. Then one can write $x = u\pi^n$, where u is a unit. The last equality gives the inclusion $\pi^n\mathfrak{O} \subsetneqq \mathfrak{a}$. Now let y be an element of $\mathfrak{O}$. We can write $y = \varepsilon\pi^m$, where ε is a unit. By our assumption $m = v(y) \geq n$, hence $y = (\varepsilon\pi^{m-n})\pi^n \in \pi^n\mathfrak{O}$, so that $\mathfrak{a} = \pi^n\mathfrak{O}$. The isomorphism results from the correspondence $a\pi^n \mapsto a \pmod{p}$. $\square$

Lemma 2.3.2. *Denoting by $|\mathfrak{O}/(\pi^e)|$ the number of elements of the quotient $\mathfrak{O}/(\pi^e)$, we have*

$$|\mathfrak{O}/(\pi^e)| = p^{ef} = p^n.$$

Proof. This follows upon combining Lemma 2.3.1 with the fact that $|\mathfrak{O}/(\pi)| = p^f$ and $|\mathfrak{O}/(\pi^e)| = |\mathfrak{O}/(\pi)| \ldots |(\pi)^{e-1}/(\pi)^e|$. $\square$

Lemma 2.3.3. *If $x \in \mathfrak{o}$ and $x \equiv 0 \pmod{\pi}$ in $\mathfrak{O}$, then $x \equiv 0 \pmod{\pi^e}$.*

Proof. The congruence $x \equiv 0 \pmod{\pi}$ is equivalent to $x = \pi x_1$, where x_1 is an element of $\mathfrak{O}$. Then we can write

$$x_1 = x_{1,0} + x_{1,1}\pi + \cdots + x_{1,e-1}\pi^{e-1},$$

where $x_{1,i} \in \mathfrak{o}$. Substituting for x_1 in $x = \pi x_1$, we obtain the identity

$$px_{1,e-1} - x + x_{1,0}\pi + \cdots + x_{1,e-2}\pi^{e-1} = 0,$$

which implies $px_{1,e-1} = x$, i.e., $x \equiv 0 \pmod{\pi^e}$. $\square$

3 Proof of Theorem 1

By the Reduction Lemma 2.1.1 we only have to show that when $s \geq em + 1$, any congruence of the form

$$a_1x_1^k + \cdots + a_sx_s^k \equiv 0 \pmod{\pi^e}, \quad a_i \in \mathfrak{O}, \qquad (5)$$

has a nontrivial solution modulo π^e. Since $\{1, \ldots, \pi^{e-1}\}$ is an $\mathfrak{o}$-basis of $\mathfrak{O}$ we can write

$$a_i = a_{i,0} + a_{i,1}\pi + \cdots + a_{i,e-1}\pi^{e-1}.$$

We look for solutions x_i only from the ring $\mathfrak{o}$. Solving (3) is equivalent to solving the system

$$\begin{array}{rcl} \sum_i^s a_{i,0}(x_1)^k & \equiv & 0 \pmod{p} \\ & \vdots & \\ \sum_i^s a_{i,e-1}(x_s)^k & \equiv & 0 \pmod{p} \end{array}$$

over $\mathfrak{o}$. Since the system consists only congruences modulo p, it suffices to solve it over the field $L(p) :=$

$\mathfrak{o}/(p)$, which is a field of characteristic p. Because the correspondence $x \mapsto x^{p^t}$ is an automorphism of $L(p)$, system (6) is equivalent to

$$\begin{pmatrix} a_{1,0} & \cdots & a_{s,0} \\ \vdots & \ddots & \vdots \\ a_{1,e-1} & \cdots & a_{s,e-1} \end{pmatrix} \begin{pmatrix} y_1^m \\ \vdots \\ y_s^m \end{pmatrix} \equiv 0 \pmod{p}, \quad (6)$$

where $x_i^{p^t}$ is replaced with y_i. By the Chevalley-Warning theorem, if $s > em$, then (6) has a nontrivial solution over $L(p)$. This completes the proof.

The result is sharp in the case $K = \mathbb{Q}_p$. Consider the equation

$$\sum_{0 \le i,j \le p-2} p^j x_{j(p-1)+i}^{p-1} = 0.$$

It is easy to apply Fermat's Little Theorem to show that this equation on $(p-1)^2$ variables has no nontrivial solution over $\mathbb{Q}_p$.

We note also that in the case of unramified K and $k = p^t$ with $p \ge 7$, Artin's conjecture (that $s \ge k^2 + 1$ ensures the solubility of (1)) follows from our theorem.

4 Proof of Theorem 2

By the same considerations as in the previous section, it suffices to show that when $s \ge p^n$, the congruence (5) has a nontrivial solution modulo π^e.

We look for solutions x_i in the set $\{0, 1\}$. Consider the sequence $0, a_1, a_1 + a_2, \dots, a_1 + a_2 + \cdots + a_s$. By virtue of $s \ge p^n$ and Lemma 2.3.2, at least two terms of this sequence are congruent modulo π^e. Let their difference be $a_u + \cdots + a_v$. Then the solution of the congruence is given by $(x_u, x_2, \dots, x_v) = (1, 1, \dots, 1)$ and $x_i = 0$ for $i < n$ or $i > v$.

We remark that the bound obtained here can be essentially improved to $s \ge k\big((1+n(p-1))^{2t+1}+1\big)-1$ by using a combinatorial theorem of John E. Olson (see J.E. Olson, A combinatorial problem on finite Abelian groups I., J. Number Theory 1 (1969), 8–10).

5 Proof of Theorem 3

By part 1 of the Reduction Lemma, we need only consider diagonal equations in which each a_i is indivisible by π; by part 2, we need only find a nontrivial solution modulo π. The Chevalley-Warning theorem solves such diagonal congruences in at least $k+1$ variables, so the Reduction Lemma shows that

$$\Gamma(k) \le k\big((k+1)-1\big)+1 = k^2+1,$$

as claimed.

6 Conclusion

When studying Skinner's work [1], we managed to develop a refinement of his method, obtaining some new upper bounds for the number of variables of a diagonal form over an extension of $\mathbb{Q}_p$. Our further goal will be to prove the general case of his claim. Let us note that the results we have obtained could be generalized to systems of diagonal equations.

7 Acknowledgments

I would like to to express my deep gratitude to my mentor, Mr. Mohsen Bahramgiri of the Massachusetts Institute of Technology, for his ongoing support and inspiring discussions. I am grateful to Prof. Hartley Rogers for providing me with such a great mentor. I am extremely thankful to Prof. Christopher Skinner for posing the problem and encouraging me to work in this field. Many thanks are due to my tutor Chris Mihelich for his constant support and professional editing of this paper. I greatly appreciate the support of the CEE, especially of the RSI staff and its director Matthew Paschke. Finally, I would like to acknowledge with thanks the help of my Bulgarian friends and colleagues Prof. Ivan Chipchakov (for enriching my knowledge about the history of the problem), Kaloyan Slavov (for sharing with me his fascination for RSI), Vesselin Dimitrov (for sending a number of valuable comments till the submission deadline), and Dr. Jenny Sendova (for helping me with some ping-pong lessons when I was stuck with the proofs).

References

[1] C.M. Skinner. Solvability of $\mathfrak{p}$-adic diagonal equations. Acta Arithmetica 75 (1996), 251–258.

[2] B.J. Birch. Diagonal equations over $\mathfrak{p}$-adic fields. Acta Arithmetica 9 (1964), 291–300.

[3] F.Q. Gouvea. *P-adic Numbers*. Springer-Verlag, Berlin (1997).

[4] S. Lang. *Algebra*. Addison-Wesley (1997).

[5] H. Davenport, D.J. Lewis. Homogeneous additive equations. Proceedings of the Royal Society of London, Series A, 274 (1964), 443–460.

[6] S. Lang. *Algebraic Number Theory.* Springer-Verlag, New York (1994).

[7] T.D. Wooley. On the local solubility of diophantine systems. Compositio Mathematica 111 (1998), 149–165.

[8] D.B. Leep and W.M. Schmidt. Systems of homogeneous equations. Inventiones Mathematicae 71 (1983), 539–549.

[9] J.P. Serre. *A Course in Arithmetic*. Springer, New York (1973).

[10] J.W. Cassels and A. Frohlich. *Algebraic Number Theory*. Academic Press, London (1967).

[11] J. Neukirch. *Algebraic Number Theory.* Springer, New York (1999).

[12] C. Chevalley. Démonstration d'une hypothèse de M. Artin. Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg (1936).

[13] L.J. Mordell. *Diophantine Equations*. Academic Press, New York (1969).

Glucocorticoid Regulation of Cell Survival in Normal and Cytokine-Treated Gastric Chief Cells

Jacob Sanders

under the direction of
Dr. Kimihito Tashima and Dr. Susan Hagen
Beth Israel Deaconess Medical Center

Abstract

Gastric chief cells secrete pepsinogen and are found at the base of gastric glands. Chief cells constitutively express the antiapoptotic protein Bcl-x_L, which may enable them to live longer than other gastric cells. We reveal that the cytokines released as part of the inflammatory response during *Helicobacter pylori* infection dose-dependently disrupt chief-cell monolayer integrity and viability. We also show that glucocorticoids dose-dependently increase expression of Bcl-x_L and hypothesize that this accounts for the ability of glucocorticoids to protect chief cells from cytokine-induced cell death. In addition, we show that glucocorticoids restrict the ability of cytokines to damage chief-cell monolayer integrity. Finally, we propose two mechanisms by which cytokines overwhelm the protective ability of glucocorticoids, resulting in chief-cell deletion and gastric atrophy during *Helicobacter pylori* infection.

1 Introduction

1.1 Gastric anatomy and chief cells

The stomach is an organ between the esophagus and the small intestine that assists in the digestion of proteins. It is divided into three regions: the cardiac region (closest to the esophagus), the pyloric region (closest to the small intestine), and the fundic region (constituting the bulk of the stomach area) [20]. The layers of tissue that line the stomach surface are collectively called the mucosa and include simple columnar epithelium, lamina propria, and muscularis mucosae. Within the fundic region, the surface is lined by mucous cells that secrete mucus. These cells also line the gastric pits. Pits are continuous with gastric glands, which contain a long neck segment (containing mostly parietal cells) and a short base segment (containing mostly chief cells) [10] (Figure 1). Parietal cells secrete HCl via an H^+/K^+-ATPase [20], and chief cells release pepsinogen from granules within their apical cytoplasm via exocytosis into the lumen of the gastric gland [10]. In the lumen, the secreted HCl activates pepsinogen to pepsin, which can hydrolyze peptide bonds to initiate protein digestion.

Chief cells develop from undifferentiated stem cells that move downward towards the base of the gastric gland. Chief cells live for approximately 190 days and are then replaced. This turnover rate is significantly longer than for other gastric cells, as parietal cells live for 54 days and surface cells for 3–4 days. Thus, chief cells must possess one or more highly specific mechanisms to regulate cell survival and cell death (apoptosis). In this study, we investigate potential regulatory mechanisms for chief-cell survival which involve glucocorticoids, Bcl-x expression, and tight junctions. We postulate that the rapid turnover of chief cells that occurs during *Helicobacter pylori* infection may be due to cytokines, which affect the regulation of tight junctions and overwhelm the ability of Bcl-x to protect against apoptosis.

1.2 Apoptosis, Bcl-x, and chief cells

Apoptosis occurs when a cell is induced to commit suicide in a controlled manner [9]. Apoptosis is important in the maintenance of tissue homeostasis, which often requires the timely and controlled death of specific cells [6]. By mediating the programmed death of cancerous cells, apoptosis also serves as a defense mechanism against tumorigenesis. An apoptotic cell undergoes cytoplasmic shrinking, DNA degradation, nuclear condensation, and mem-

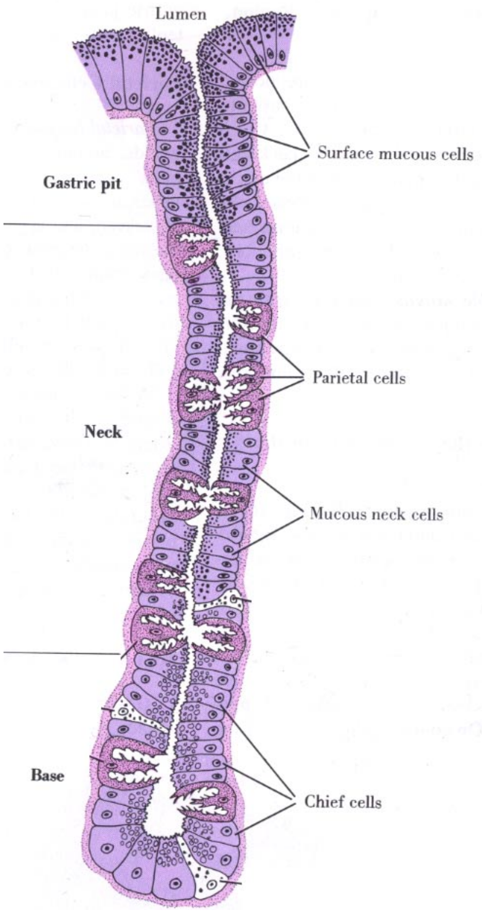


Figure 1: This diagram [20] shows the location of the different cells in the gastric mucosa. Chief cells are found at the base of gastric glands.

brane blebbing [1], in addition to activation of cell-death proteins and pathways that are cell-type specific.

Apoptosis is regulated through complex signaling pathways, which include proapoptotic and antiapoptotic (prosurvival) proteins [6] from the Bcl-2 family of proteins. Proteins in the Bcl-2 family possess at least one of four conserved motifs known as Bcl-2 homology (BH) domains [1]. Antiapoptotic proteins in this family generally contain at least BH1 and BH2, whereas many (but not all) proapoptotic proteins lack these domains. Different tissues express different members of the Bcl-2 protein family, but the ratio of proapoptotic to antiapoptotic proteins appears to determine whether or not a cell undergoes apoptosis [1]. The mechanism through which these proteins control apoptosis is not completely understood but appears in certain cases to involve heterodimerization between proapoptotic and antiapoptotic proteins [1].

Little is known about the expression of Bcl-2 family proteins in the stomach, in general, or in chief cells. Neu *et al.* [17] found that rat gastric chief cells express the prosurvival protein Bcl-2 and the proapoptotic protein Bax, a finding that was confirmed by Hagen *et al.* (unpublished data). Recently, rat chief cells were also shown to express Bcl-x, a second survival protein that is usually expressed in long-lived cells.

The *bcl-x* gene undergoes alternative splicing, leading to two ubiquitous isoforms: Bcl-x_L and Bcl-x_S [2]. (Three other isoforms have been identified, but are not ubiquitous.) Bcl-x_L is a prosurvival protein that, like Bcl-2, contains the BH1 to BH4 domains [1]. Bcl-x_S contains a deletion of 63 amino acid residues relative to Bcl-x_L and contains only BH3 and BH4 domains [2]. Minn *et al.* [14] showed that Bcl-x_S is proapoptotic and can antagonize Bcl-x_L. This is consistent with the observation that the deleted residues in Bcl-x_S are those with greatest similarity to Bcl-2 [2]. Pecci *et al.* [18] identified five promoters (P1 to P5) of the *bcl-x* gene in mice and reported that tissue-specific promoter choice influences alternative splicing and, thus, the balance of Bcl-x isoforms. Promoter choice in chief cells may affect the ratio of Bcl-x_L to Bcl-x_S and regulate the balance between apoptosis and survival.

1.3 Tight junctions and chief cells

Tight junctions are structures that seal adjacent cells together [21]. They are present in epithelial tissues, including the simple columnar epithelium that lines the gastric lumen. Tight junctions are important in gastric glands because they prevent mucus, HCl, and pepsinogen from leaking out of the stomach lumen. Tight junctions also selectively prevent ions and molecules, especially secreted H^+, from entering gastric cells from the stomach lumen. The apical membrane of gastric cells, which faces the lumen, is generally impermeable to these molecules. However, without tight junctions, these molecules would be able to pass between gastric cells and then enter gastric cells via the basolateral membrane, which is more permeable. Thus, tight junctions are important in the gastric mucosa because they can protect gastric cells, including chief cells, from ions and molecules that would affect their function and survival.

Tight junctions are composed of a series of transmembrane proteins. Two cells forming a tight junction possess proteins called claudins in their plasma membranes. The claudins in the membranes of the two cells bind to each other, thus sealing the two cells together [21]. Occludin is also found in the membranes of two cells forming a tight junction, but its role in tight-junction formation is less clearly defined [21]. Finally, ZO proteins anchor claudins and occludin to the cytoskeleton of the cells forming a tight junction [21].

Transepithelial electrical resistance (TER) is commonly used to verify that tight junctions are present in a monolayer of epithelial cells. A high TER is associated with a tight monolayer that has low ion permeability. Conversely, a low TER indicates a leaky monolayer with high permeability. Thus, measurement of TER is used in experiments to determine whether a treatment affects monolayer integrity and permeability [21].

1.4 Glucocorticoid regulation of chief-cell survival

Glucocorticoids are steroid hormones implicated in the regulation of the *bcl-x* gene. Viegas *et al.* [26] reported that treatment of cells with dexamethasone (a synthetic glucocorticoid) allowed glucocorticoid receptor to bind to two hormone response element (HRE)-like sequences on the distal promoter P4 of *bcl-x*. This resulted in the expression of the antiapoptotic Bcl-x_L isoform. Another study by Chang *et al.* [4] showed that dexamethasone suppressed apoptosis in the human gastric cancer TMK-1 cell line through modulation of *bcl-x* expression. Inducing apoptosis in the cells led to increased expression of Bcl-x_S and cell death. In contrast, treatment with dexamethasone decreased Bcl-x_S expression and increased Bcl-x_L expression to favor cell survival.

Glucocorticoids are also implicated in the regulation of tight junctions. A study by Zettl *et al.* [27] found that glucocorticoids induce mouse mammary epithelial cells to form tight junctions by modulating the expression of tight-junction proteins in cells in a similar manner as described for *bcl-x*.

Because *bcl-x* is expressed at a high level in gastric chief cells and because adjacent chief cells form tight junctions, we postulate that glucocorticoids might be involved in the control of chief-cell survival by regulating *bcl-x* expression and/or monolayer integrity. Tseng *et al.* [24] reported that adrenalectomized rats exhibit a dramatic decrease in the number of mature gastric chief cells, suggesting that glucocorticoids may regulate the survival of chief cells.

We hypothesize that glucocorticoids protect chief cells from cell death, in part, by increasing the expression of the antiapoptotic Bcl-x_L isoform and by preserving monolayer integrity.

1.5 *Helicobacter pylori*, cytokines, and chief-cell death

Helicobacter pylori (HP) is a gram-negative rod-shaped bacterium that infects the stomach of more than half of the global population [16]. It survives the acidic environment of the stomach by catalyzing the breakdown of urea into the basic bicarbonate and ammonia species [13]. HP colonizes the gastric mucosa and causes gastritis, followed by atrophy, metaplasia, and cancer. Atrophy, defined as the deletion of parietal and chief cells in gastric glands, is the major factor that predisposes to gastric cancer progression during HP infection [13].

Although specific pathways remain to be defined, infection with HP triggers an inflammatory response in the gastric mucosa [22]. This is associated with the release of proinflammatory cytokines from T helper-1 (T_H1) cells, including tumor necrosis factor (TNF)-α, interleukin (IL)-1β, and interferon (IFN)-γ [8, 22]. These cytokines attract leukocytes to the site of infection. In addition, IL-1β and TNF-α inhibit the secretion of gastric acid by parietal cells, allowing HP to further colonize the gastric mucosa [8]. In the fundus, the increase in inflammation is associated with increased apoptosis of gastric epithelial cells, resulting in gastric atrophy. In cells of the intestinal mucosa, cytokines released during inflammatory bowel disease are known to alter tight-junction structure, increasing their permeability and making them leaky [3, 5].

In our experiments, we use $\mathrm{T_H}1$ cytokines (TNF-α, IL-1β, and IFN-γ) to simulate the cytokines released by inflammatory cells during HP infection. We hypothesize that this combination of cytokines causes chief-cell death. Our experiments aim to confirm this hypothesis, and to shed light on the mechanisms by which cytokines kill chief cells. One possibility is that the cytokines stimulate apoptosis of chief cells by upregulating the proapoptotic Bcl-$\mathrm{x_S}$ isoform or by downregulating the antiapoptotic Bcl-$\mathrm{x_L}$ isoform. Alternately, the cytokines may disrupt the integrity of the chief-cell monolayer, possibly by altering the expression of tight-junction proteins. Another possibility is that the cytokines kill chief cells through some combination of these two mechanisms. For example, a change in Bcl-x protein production may result in disruption of the chief-cell monolayer, ultimately causing anoikis (detachment-induced apoptosis). Because Bcl-x expression and monolayer integrity are regulated by glucocorticoids, we also investigate the ability of dexamethasone (a synthetic glucocorticoid) to protect chief cells from cell death induced by cytokines.

2 Materials and Methods

2.1 Chief-cell isolation and purification

Stomachs were surgically removed under heavy anesthesia from nonfasted rats weighing between 180 and 200 g. The pyloric region was cut, and the stomach was everted to expose the epithelium. After ligating both ends of the stomach and filling with protease solution, successive fractions of cells were collected every 30 minutes from buffer solutions that covered the epithelium and were maintained at 37°C in a shaking water bath. Early fractions containing surface cells, neck cells, and parietal cells were discarded. The final fraction contained mostly chief cells and was kept for further purification.

This chief-cell fraction was pelleted by centrifugation, resuspended in buffer, filtered, and then subjected to centrifugal elutriation at 40 ml/minute to eliminate the non–chief-cell population. The elutriated chief-cell fraction was then further purified by density-gradient centrifugation. For this, the suspension of chief cells was transferred to tubes that contained a linear 40–80% density gradient of Accudenz (Accurate Chemical, Westbury, NY) and centrifuged at 500×g for 10 minutes. Because chief cells have the highest density of all gastric cells, they were collected from the bottom of the linear density gradient tubes, washed in buffer, pelleted by centrifugation, and resuspended in a 1:1 mixture of Ham's F12 and Dulbecco's Modified Eagle medium supplemented with heat-inactivated 10% fetal bovine serum, antibiotics, insulin, and hepatocyte growth factor. Cells were counted using a hemocytometer, and 1.6×10^6 cells/ml were plated. Hydrocortisone or dexamethasone, both commercially available glucocorticoids, were also added in some studies as described below.

2.2 Chief-cell culture and treatment

Transwell plates and 6-cm circular dishes, both coated with type-1 collagen, were used to culture cells. Cells were grown for 96 hours to achieve a confluent monolayer; cell culture medium was replaced on a regular basis. Three experiments were done with cultured chief cells.

2.2.1 Glucocorticoids and chief-cell survival

To assess the effect of glucocorticoids on cell viability, barrier properties, and Bcl-x protein production in the absence of cytokine challenge, chief-cell cultures were given the following glucocorticoid treatments: (a) no glucocorticoid; (b) 10^{-9} M dexamethasone; (c) 10^{-6} M dexamethasone; (d) 1 μg/ml hydrocortisone.

2.2.2 Chief-cell survival in the presence of $\mathrm{T_H}1$ cytokines

Using an *in vitro* assay to assess the effect of inflammation during HP infection on cell viability, barrier properties, and Bcl-x protein production, chief-cell cultures were treated with or without a cytokine mix (cytomix) containing 10 ng/ml each of TNF-α, IL-1β, and IFN-γ. Dose dependence was tested by varying the percentage of the cytomix used to treat cell cultures: (a) no cytomix; (b) 10% cytomix; (c) 30% cytomix; (d) 100% cytomix.

2.2.3 Chief-cell survival with cytokines in the presence of glucocorticoids

To investigate the ability of glucocorticoids to protect against chief-cell death induced by the $\mathrm{T_H}1$ cytomix, chief-cell cultures were treated in a dose-dependent manner both with the cytomix and with dexamethasone in the following proportions: (a) 10^{-9} M dexamethasone and 10% cytomix; (b) 10^{-9} M dexamethasone and 30% cytomix; (c) 10^{-9} M dexamethasone and 100% cytomix; (d) 10^{-6} M dexamethasone and 10% cytomix; (e) 10^{-6} M dexamethasone and 30% cytomix; (f) 10^{-6} M dexamethasone and 100% cytomix.

To investigate the effects of dexamethasone and cytokine treatments on the monolayer integrity and barrier properties of chief cells, the transepithelial electrical resistance (TER) was measured in the Transwell plates at 0, 6, 12, 24, 30, 36, and 48 hours following each treatment. After 48 hours of treatment, the cell cultures in the Transwell plates were used for the cell viability assay.

2.3 Cell-viability assay

The cell-viability assay was used to determine the relative numbers of living, viable cells following the glucocorticoid and cytokine treatments. For each group of treated cells, the cell culture medium in Transwell plates was aspirated. The cells were washed with phosphate-buffered saline and fixed with ice-cold 100% methanol at 4°C for 15 minutes. The cells were stained with 0.1% crystal violet solution, incubated at room temperature for 5 minutes, and washed with warm tap water.

After drying overnight, Transwells were then treated with 0.5% sodium dodecyl sulfate (SDS) solution and agitated for 30 minutes at room temperature to facilitate cell lysis and elaboration of the stain into solution. Lysates were transferred to wells on a 96-well plate and diluted with SDS, and the absorbances at 590 nm and 650 nm were determined by spectroscopy. Stain intensity, and thus cell viability, was determined by subtracting the absorbance readings.

2.4 SDS-PAGE and Western blotting

Using the 6-cm dishes of cultured chief cells, sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and Western blotting were used to determine the relative expression of Bcl-x following glucocorticoid and cytokine treatments, as described in Section 2.2. Chief cells were washed and treated with radioimmunoprecipitation assay (RIPA) buffer containing proteolytic inhibitors to enable cell lysis and protein solubilization while preventing degradation of individual polypeptides. The cells were scraped, sonicated to disrupt cell integrity, and cleared of cell debris by centrifugation. The supernatant was collected, and overall protein concentration of each sample was determined using the bicinchoninic acid (BCA) protein-assay kit (Pierce Biotechnology, Rockford, IL).

Protein samples were separated by SDS-PAGE. Equal amounts of each protein sample (10 μg) were mixed with gel buffer containing 2-mercaptoethanol, which reduces (breaks) disulfide bonds. The samples were loaded onto a polyacrylamide gel (12.5% base and 4.0% stack) and allowed to run for 35 minutes with a current of 200 A.

The resulting gels were electroblotted onto nitrocellulose membranes for 60 minutes at 15 V, and the presence of protein was confirmed by Ponceau staining. Each blot was blocked overnight at 4°C in a solution containing 5% nonfat dry milk to block nonspecific binding sites. Blots were incubated for 2 hours at room temperature with the primary antibody in 1% nonfat milk solution. Blots were then washed and incubated for 1 hour at room temperature with the secondary antibody in 1% nonfat milk solution. The primary antibody used was rabbit anti-Bcl-$x_{S/L}$ L-19 (1:500) (Santa Cruz Biotechnology, Santa Cruz, CA). The secondary antibody used was goat anti-rabbit horseradish peroxidase (HRP)-conjugated antibody (1:2000). Blots were developed using the Enhanced Chemiluminescence (ECL) Kit (Amersham, Arlington Heights, IL), and the bands visualized by autoradiography using Blue Bio Film (Denville Scientific, South Plainfield, NJ).

3 Results

3.1 Glucocorticoids upregulate Bcl-x_L expression in cultured chief cells

The synthetic glucocorticoid dexamethasone increased the expression of the antiapoptotic Bcl-x_L isoform in chief cells (Figure 2). As the gel shows in lane 1, cultured chief cells express Bcl-x_L constitutively after withdrawal of glucocorticoid treatment for 48 hours. Thus, the fetal bovine serum used to culture the chief cells maintains the constitutive expression of Bcl-x_L for at least 48 hours. However, the addition of dexamethasone to chief cell cultures significantly increased the expression of Bcl-x_L in a dose-dependent manner, as shown in lanes 2 and 3. Bcl-x_L expression in chief cells increased significantly with 10^{-9} M dexamethasone (lane 2), and increased even further with 10^{-6} M dexamethasone (lane 3). Chief cells did not constitutively express the proapoptotic Bcl-x$_S$ isoform, and glucocorticoids did not stimulate Bcl-x$_S$ expression (Figure 2).

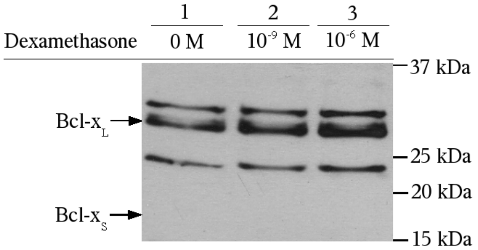


Figure 2: This Western blot shows that Bcl-x$_L$ expression increases as glucocorticoid concentration increases. The thick band is Bcl-x$_L$, while the two other thin bands represent contamination. Bcl-x$_S$ is not expressed in any lane.

3.2 Chief-cell viability and TER do not depend on glucocorticoid concentration in the absence of cellular stress

Dexamethasone treatment, which upregulated Bcl-x$_L$ expression (Figure 2), did not affect chief-cell viability over 48 hours (Figure 3). Hydrocortisone (1 μg/ml) was used as a control because it is originally added to cultured chief cells to allow them to reach confluence and maturity prior to the differential 48-hour glucocorticoid treatments. As Figure 3 shows, treatment with dexamethasone (known to be about 40 times as potent as hydrocortisone) did not increase cell viability relative to treatment with hydrocortisone. Thus, in the absence of cellular stress, glucocorticoids with high potency and present in high concentrations were not required to keep chief cells alive.

In the absence of cellular stress, glucocorticoids did not affect TER to a large extent (Figure 4). From 24 to 48 hours following treatment, no changes in chief cell TER greater than 500 $\Omega\cdot$cm^2 were found for any glucocorticoid treatment (either hydrocortisone or dexamethasone). These results indicate that, in the absence of cellular stress, integrity of the chief cell monolayer was preserved for at least 48 hours regardless of glucocorticoid potency and concentration.

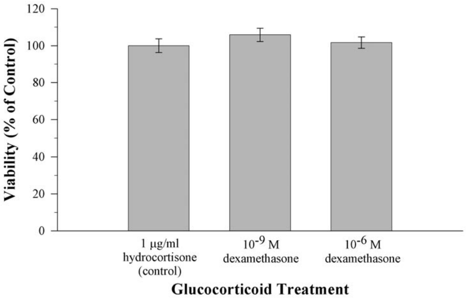


Figure 3: This graph shows chief cell viability in response to glucocorticoid treatments. Relative to hydrocortisone, dexamethasone (which is 40 times as potent) did not affect chief cell viability. Error bars in this and all subsequent graphs show standard error.

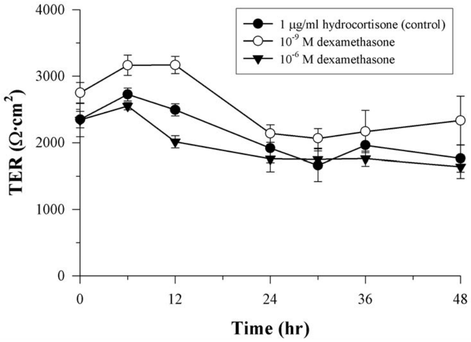


Figure 4: This graph shows the effect of different glucocorticoid treatments on TER. After an initial increase in TER in all groups, dexamethasone at 10^{-9} M or 10^{-6} M did not affect TER.

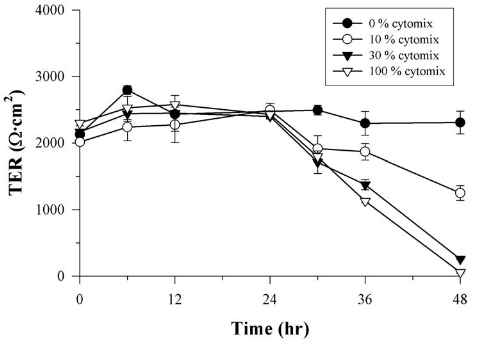


Figure 5: This graph shows the effect of increasing cytomix concentrations on chief-cell TER. The cytomix decreased TER in a dose-dependent manner by 48 hours after addition to chief-cell cultures.

3.3 T_H1 cytokines reduce chief-cell TER and viability

The cytomix was shown to decrease both chief-cell TER and chief-cell viability in a dose-dependent manner (Figures 5 and 6). When TER was measured in control chief cells that were not treated with the cytomix, TER remained stable over 48 hours. Chief cells treated with the cytomix maintained a relatively stable TER for 24 hours following treatment. However, after 48 hours, there was a dose-dependent decrease in TER. Thus, after 48 hours, our data suggests that cytokines disrupted the barrier properties of the chief-cell cultures.

To determine whether the reduction in TER correlates with changes in viability, we measured chief-cell viability in response to the cytomix and showed that the cytomix decreased chief-cell viability in a dose-dependent manner (Figure 6). After treatment for 48 hours, the number of viable chief cells was reduced by 20% with 10% cytomix, 40% with 30% cytomix, and 50% with 100% cytomix. Although there is correlation between TER and survival data, it was of interest that TER was affected far more than viability after 48 hours of treatment (compare Figures 5 and 6). For instance, TER was reduced nearly to 0 with the 100% cytomix, but viability was only reduced by 50%.

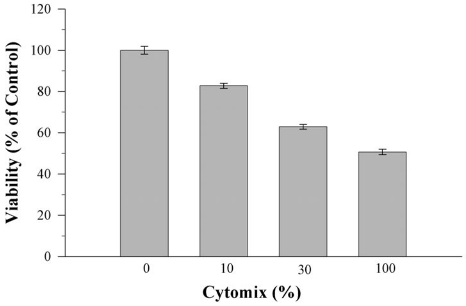


Figure 6: This graph shows the effect of increasing cytomix concentrations on chief-cell viability. The cytomix decreased viability in a dose-dependent manner.

3.4 Glucocorticoids protect chief cells from the effects of cytokines

When chief cells were treated with both dexamethasone and T_H1 cytokines, dexamethasone was able to protect chief cells from the effects of the cytokines on both TER and viability (Figures 7 and 8). Shown in Figure 7, both 10^{-9} M (A) and 10^{-6} M (B) dexamethasone were able to prevent reduction in TER over 48 hours in chief cells treated with all dilutions of the cytomix. Only in chief cells treated with 100% cytomix was dexamethasone unable to completely prevent the decrease in TER (Figure 7). These results indicate that glucocorticoids can effectively prevent a cytokine-induced decrease in TER.

Similar to results with TER, dexamethasone also protected chief cells from cytokine-induced death (Figure 8). However, neither 10^{-9} M (A) nor 10^{-6} M (B) dexamethasone were able to exert complete protection against the cytomix. Even with dexamethasone present, the 100% cytomix decreased the number of viable chief cells by about 30% (relative to chief cells not treated with cytokines). However, this is less than the 50% decrease in chief-cell viability caused by the 100% cytomix in the absence of dexamethasone.

4 Discussion

4.1 Glucocorticoids regulate Bcl-x_L expression in chief cells

Our results showed that dexamethasone, a synthetic glucocorticoid, increased Bcl-x_L expression in chief cells in a dose-dependent manner. We propose that

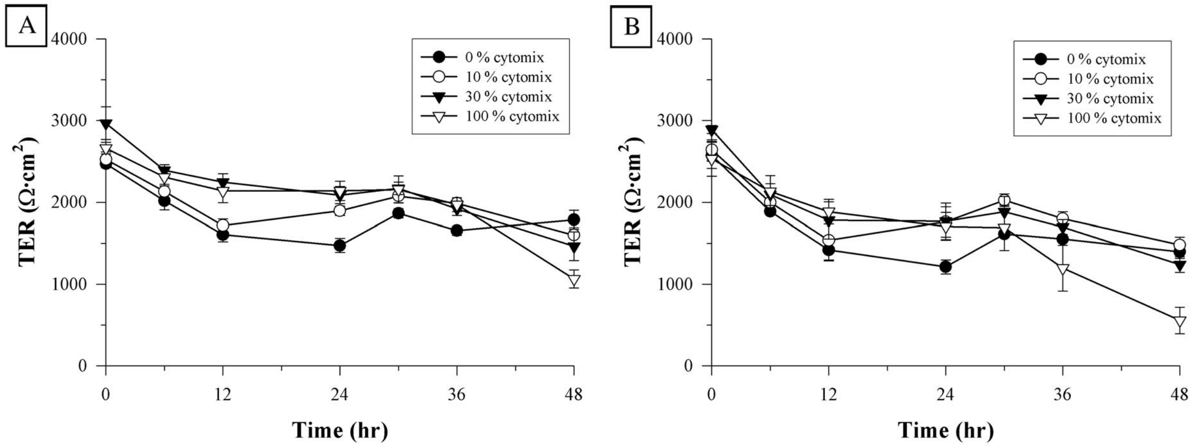


Figure 7: These graphs reveal the ability of 10^{-9} M (A) and 10^{-6} M (B) dexamethasone to protect chief cells from cytokine-induced decrease in TER.

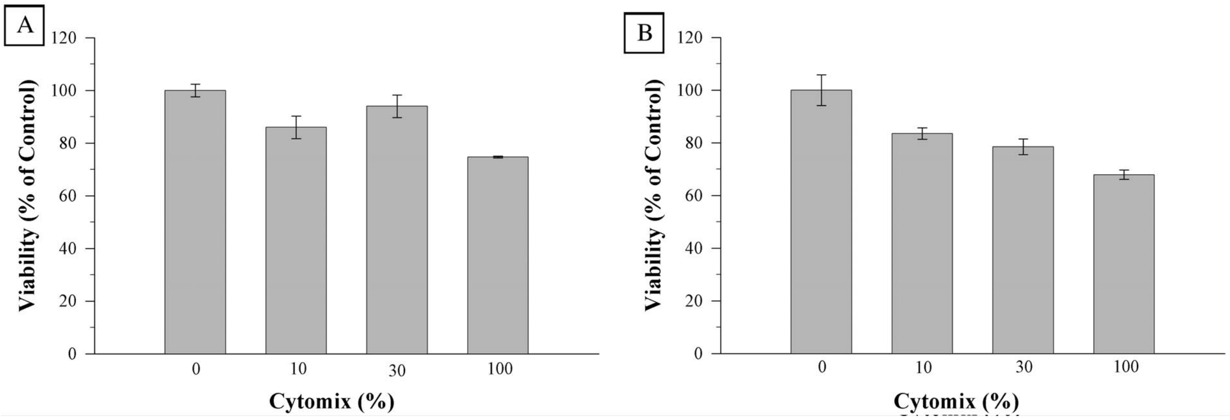


Figure 8: These graphs reveal the ability of 10^{-9} M (A) and 10^{-6} M (B) dexamethasone to preserve cell viability even in the presence of cytokines.

glucocorticoids bind to the intracellular glucocorticoid receptor (GR) in chief cells, which acts as a transcription factor by binding to hormone response element (HRE)-like sequences on one of the five promoters of the *bcl-x* gene. Because Viegas *et al.* [26] identified two HRE-like sequences on the extended P4 promoter, we hypothesize that chief cells use the P4 promoter to induce expression of Bcl-x_L.

That chief cells respond to glucocorticoids suggests the presence of a GR receptor. However, Kanemasa *et al.* [11] showed that GR receptors are found only in parietal cells in the gastric mucosa. This result is confusing in light of our present data and data showing that adrenalectomy, or reduction of glucocorticoids, significantly impacts chief-cell development [24]. Further studies using RU 38486, a GR receptor antagonist, would be required to prove that the response we see to endogenous dexamethasone is specific and acting at GR receptor to increase Bcl-x_L expression in chief cells.

4.2 Chief cells do not require glucocorticoids to survive in absence of cellular stress

Our results showed that, in the presence of serum, addition of glucocorticoid is not required to maintain chief-cell survival or TER for 48 hours in culture. We show that chief cells constitutively express Bcl-x_L and, even when glucocorticoids are added to increase chief-cell production of Bcl-x_L, chief-cell survival is not altered. Thus, constitutively-expressed Bcl-x_L is sufficient to maintain chief-cell survival for 48 hours in the absence of cellular stress. Consistent with our results, the findings of Boise *et al.* [2] show that increased Bcl-x_L expression is important only in the face of cellular stress.

4.3 Cytokines kill chief cells: possible mechanisms

Our data clearly show that T_H1 cytokines decrease chief-cell TER and viability in a dose-dependent manner. In this respect, chief cells are unique in comparison to other cells of the gastrointestinal tract. For example, Clayburgh *et al.* [5] found that T_H1 cytokines affected tight junctions between intestinal cells, resulting in barrier defects, but did not reduce the viability of epithelial cells.

One possible mechanism for cytokine-induced chief-cell death involves the upregulation of Bcl-x_S by the NF-κB protein family. NF-κB proteins are known to translocate to the nucleus in response to cytokine-induced stress [23]. NF-κB proteins function as transcription factors and regulate the *bcl-x* gene through an NF-κB binding sequence on the *bcl-x* promoter [12]. While some studies [25] have shown that NF-κB increases production of Bcl-x_L, others [7] have implicated NF-κB in the upregulation of Bcl-x_S. Thus, we propose that cytokines cause NF-κB proteins to localize to the nucleus of chief cells, where they increase production of Bcl-x_S, which kills the cells. This mechanism requires further investigation: it remains to be shown that cytokines result in localization of NF-κB proteins to the nucleus and in increased expression of Bcl-x_S.

4.4 Glucocorticoids protect chief cells from cellular stress

In the absence of cellular stress, glucocorticoids had little effect on chief-cell barrier properties or viability. However, in the presence of cytokine-induced cellular stress, glucocorticoids preserved the chief-cell barrier and kept chief cells alive.

As previously discussed, glucocorticoids increase Bcl-x_L expression in chief cells in a dose-dependent manner. While this increase in Bcl-x_L was not required for chief-cell survival in the absence of cellular stress, we propose that the increase in Bcl-x_L is required for chief-cell survival in the presence of cellular stress (such as T_H1 cytokines). Thus, the glucocorticoid-regulated increase in Bcl-x_L expression may explain the ability of glucocorticoids to protect chief cells from cytokine-induced death.

Furthermore, glucocorticoids are known to upregulate tight-junction proteins such as occludin. While this upregulation of occludin is not necessary to preserve chief-cell tight junctions in the absence of cellular stress, it may be necessary to protect chief-cell tight junctions and barrier properties from the damaging effects of cytokines.

Of course, Bcl-x_L and occludin are not the only proteins regulated by glucocorticoids. It is possible that glucocorticoids also upregulate other protective proteins which help protect chief cells from the stress of cytokines.

4.5 Cytokines overwhelm glucocorticoids in *H. pylori* infection

HP infection triggers an inflammatory response leading to the release of proinflammatory T_H1 cytokines. Nonetheless, glucocorticoids are constitutively secreted by the adrenal glands. Based on our results, we might expect these glucocorticoids to protect chief cells from cytokine-induced cell death. However, this is clearly not the case, as it is known that chief cells are deleted during HP infection. We propose two possible mechanisms by which cytokines may overwhelm the ability of glucocorticoids to protect chief cells during HP infection.

The first mechanism is simple: the cytokines released during HP infection may be present in a concentration that is too high for the protective effects of glucocorticoids to be significant. This mechanism is consistent with our results, which show that neither 10^{-9} M dexamethasone nor 10^{-6} M dexamethasone were able to completely protect chief cells when exposed to the 100% cytokine mix. Thus, glucocorticoids cannot completely protect against high concentrations of cytokines. As a result, chief cells are deleted in large numbers during HP infection.

As a second potential mechanism, HP infection may interfere with a step in the pathway by which glucocorticoids protect chief cells from tight-junction damage and cell death. A study by Peltier *et al.* [19] found that transforming growth factor (TGF)-β1, via Smad 2/3 and AP1 transcription factor activation, increases the expression of intracellular glucocorticoid receptor in inflammatory cells. Interestingly, Monteleone *et al.* [15] recently showed that HP, via IFN-γ, induced Smad 7 and significantly reduced Smad 3 activity. As a result, chief-cell expression of glucocorticoid receptor would decrease, and this would decrease chief-cell responsiveness to glucocorticoids. Thus, glucocorticoids would not be able to increase production of the antiapoptotic Bcl-x_L protein. Nonetheless, the proapoptotic Bcl-x_S protein might still be produced via the glucocorticoid-independent pathway involving cytokines and NF-κB. In the absence of glucocorticoid receptor, glucocorticoids would be unable to protect chief cells from tight-junction damage and cell death.

While these two mechanisms are very different and highly speculative, they both explain how constitutively expressed glucocorticoids may be unable to protect chief cells from cytokines during HP infection.

5 Conclusion

In this paper, we demonstrated that cytokines kill chief cells. We also showed that glucocorticoids can protect chief cells from cytokine-induced death. We identified glucocorticoid-induced upregulation of Bcl-x_L expression and preservation of monolayer integrity as two possible protective mechanisms. Finally, we proposed two mechanisms to explain how the ability of glucocorticoids to protect chief cells from cytokines may be overwhelmed during HP infection.

6 Acknowledgments

I am greatly indebted to my mentor, Dr. Susan Hagen of Beth Israel Deaconess Medical Center. Dr. Hagen provided me with background reading, discussed my project with me, allowed me to perform experiments in her laboratory, helped me interpret my results, and edited this paper. I also thank Dr. Kimihito Tashima, a postdoctoral fellow in the laboratory who spent countless hours helping me master experimental techniques, organizing my project, and gathering the TER data presented in this paper. I also express gratitude to Dr. Asli Muvaffak, another postdoctoral fellow who was always ready to answer any of my immediate questions.

I am also grateful for the assistance provided by my tutor, Viviana Risca, in discussing the progress of my project and providing advice for this paper. Finally, I thank the Research Science Institute and the Center for Excellence in Education for giving me the opportunity to perform research in such a wonderful setting.

References

[1] J.M. Adams and S. Cory. The Bcl-2 protein family: arbiters of cell survival. Science 281 (1998), 1322–1326.

[2] L.H. Boise, M. Gonzalez-Garcia, C.E. Postema, L. Ding, T. Lindsten, L.A. Turka, X. Mao, G. Nunez, and C.B. Thompson. *bcl-x*, a *bcl-2*-related gene that functions as a dominant regulator of apoptotic cell death. Cell 74 (1993), 597–608.

[3] M. Bruewer, A. Luegering, T. Kucharzik, C.A. Parkos, J.L. Madara, A.M. Hopkins, and A. Nusrat. Proinflammatory cytokines disrupt epithelial barrier function by apoptosis-independent mechanisms. Journal of Immunology 171 (2003), 6164–6172.

[4] T. Chang, M. Hung, S. Jiang, J. Chu, L. Chu, and L. Tsai. Dexamethasone suppresses apoptosis in a human gastric cancer cell line through modulation of *bcl-x* gene expression. Federation of European Biochemical Sciences 415 (1997), 11–15.

[5] D.R. Clayburgh, L. Shen, and J.R. Turner. A porous defense: the leaky epithelial barrier in intestinal defense. Laboratory Investigation (2004), 1–10.

[6] N.N. Danial and S.J. Korsmeyer. Cell death: critical control points. Cell 116 (2004), 205–219.

[7] J. Doostzadeh-Cizeron, S. Yin, and D. Goodrich. Apoptosis induced by the nuclear death domain protein p84N5 is associated with caspase-6 and NF-κB activation. Journal of Biological Chemistry 275 (2000), no. 33, 25336–25341.

[8] E.M. El-Omar. The importance of interleukin 1β in *Helicobacter pylori* associated disease. Gut 48 (2001), 743–747.

[9] D.R. Green and J.C. Reed. Mitochondria and apoptosis. Science 281 (1998), 1309–1312.

[10] S. Ito. Functional gastric morphology. In: L.R. Johnson, editor. *Physiology of the Gastrointestinal Tract*. 2nd ed. Raven Press, New York, NY (1987), 817–851.

[11] H. Kanemasa, H. Ozawa, H. Konishi, T. Ito, M. Nishi, S. Mitsufuji, T. Kodama, T. Hattori, and M. Kawata. Distribution of glucocorticoid receptor immunoreactivity in gastric mucosa of normal and adrenalectomized rats. Digestive Diseases and Sciences 44 (1999), no. 10, 2081–2087.

[12] H. Lee, H. Dadgostar, Q. Cheng, J. Shu, and G. Cheng. NF-κB-mediated up-regulation of Bcl-x and Bfl-1/A1 is required for CD40 survival signaling in B lymphocytes. Proceedings of the National Academy of Sciences 96 (1999), 9136–9141.

[13] B. Marshall. *Helicobacter pylori*: 20 years on. Clinical Medicine 2 (2002), no. 2, 147–152.

[14] A.J. Minn, L.H. Boise, and C.B. Thompson. Bcl-x_S antagonizes the protective effects of Bcl-x_L. Journal of Biological Chemistry 271 (1996), no. 11, 6306–6312.

[15] G. Monteleone, G. Del Vecchio Blanco, G. Palmieri, P. Vavassori, I. Monteleone, A. Colantoni, S. Battista, L.G. Spagnoli, M. Romano, M. Borrelli, T.T. MacDonald, and F. Pallone. Induction and regulation of Smad7 in the gastric mucosa of patients with *Helicobacter pylori* infection. Gastroenterology 126 (2004), no. 3, 674–682.

[16] M. Naumann and J.E. Crabtree. *Helicobacter pylori*-induced epithelial cell signalling in gastric carcinogenesis. Trends in Microbiology 12 (2004), no. 1, 29–36.

[17] B. Neu, H. Herrmuth, F. Ernst, W. Vaupel, W. Reindl, P. Hutzler, M. Atkinson, M. Classen, and W. Schepp. Differential expression of CD95, Bcl-2, and Bax in rat gastric chief and parietal cells. Microscopy Research and Technique 53 (2001), 377–388.

[18] A. Pecci, L. R. Viegas, J.L. Baranao, and M. Beato. Promoter choice influences alternative splicing and determines the balance of isoforms expressed from the mouse *bcl-X* gene. Journal of Biological Chemistry 276 (2001), no. 24, 21062–21069.

[19] J. Peltier, J. Perez, A. Bellocq, B. Escoubet, B. Fouqueray, and L. Baud. Transforming growth factor-beta 1 increases glucocorticoid binding and signaling in macrophages through a Smad- and activated protein-1-mediated process. Kidney International 63 (2003), no. 6, 2028–2036.

[20] M.H. Ross, L.J. Romrell, and G.I. Kaye. *Histology: A Text and Atlas*. 3rd ed. William & Wilkins, Baltimore, MD (1995).

[21] E.E. Schneeberger and R.D. Lynch. The tight junction: a multifunctional complex. American Journal of Physiology: Cell Physiology 286 (2003), 1213–1228.

[22] A.R. Sepulveda and D.Y. Graham. Role of *Helicobacter pylori* in gastric carcinogenesis. Gastroenterology Clinics 31 (2002), no. 2, 517–535.

[23] L. Sevilla, A. Zaldumbide, P. Pognonec, and K.E. Boulukos. Transcriptional regulation of the *bcl-x* gene encoding the anti-apoptotic Bcl-x_L protein by Ets, Rel/NFκB, STAT and AP1 transcription factor families. Histology and Histopathology 16 (2001), 595–601.

[24] C. Tseng, K.L. Schmidt, and L.R. Johnson. Hormonal effects on development of the secretory apparatus of chief cells. American Journal of Physiology: Gastrointestinal Liver Physiology 253 (1987), 274–283.

[25] T. Tsukahara, M. Kannagi, T. Ohashi, H. Kato, M. Arai, G. Nunez, Y. Iwanaga, N. Yamamoto, K. Ohtani, M. Nakamura, and M. Fugii. Induction of Bcl-x_L expression by human T-cell leukemia virus type 1 Tax through NF-κB in apoptosis-resistant T-Cell transfectants with Tax. Journal of Virology 73 (1999), no. 10, 7981–7987.

[26] L.R. Viegas, G.P. Vicent, J.L. Baranao, M. Beato, and A. Pecci. Steroid hormones induce *bcl-X* gene expression through direct activation of distal promoter P4. Journal of Biological Chemistry 279 (2004), no. 11, 9831–9839.

[27] K.S. Zettl, M.D. Sjaastad, P.M. Riskin, G. Parry, T.E. Machen, and G.L. Firestone. Glucocorticoid-induced formation of tight junctions in mouse mammary epithelial cells *in vitro*.

Proceedings of the National Academy of Sciences
89 (1992), 9069–9073.

Abstracts of Award-Winning Oral Presentations

Using Algebraic Coding for Transferring Files in Peer-to-Peer Networks

Szymon Acedański Under the direction of
Maria Sklodowska-Curie HS Prof. Muriel Médard
Katowice, Poland MIT LIDS

Algebraic coding is a method of improving capacity of multicast networks. Inspired by the mathematical ideas behind this method, researchers at MIT suggested a new method of transferring files in peer-to-peer (unicast) networks, called *random linear coding storage*. In this paper we simulate the proposed method. The analysis carried out shows, that it is both more efficient and more reliable than the simple *random uncoded storage* approach.

Characterization of Co(II) Binding to PAMAM G4-NH_2 Dendrimer in Aqueous Solutions Using UV and EPR Spectroscopy

Khaled Al-Rabe Under the direction of
Al-Bayan Bilingual School Dr. Mamadou S. Diallo
Khaldiya, Kuwait Caltech

Preliminary experiments have shown that generation-4 polyamidoamine (PAMAM) dendrimer (G4-NH_2) is a high-capacity chelating agent for Co(II) in aqueous solutions. However the structures of the Co(II)-dendrimer complexes are not known. This paper describes an experimental investigation of Co(II) binding to a G4-NH_2 PAMAM dendrimer in aqueous solutions using UV-VIS and EPR spectroscopy. The overall results suggest that the Co(II)-dendrimer complexes include four dendrimer tertiary amine groups in a distorted octahedral shape.

Investigation of Graphs Having Identical Path Layer Matrices

Hannah Chung Under the direction of
LBJ HS Mr. Alan Leung
Austin, TX MIT

The *path layer matrix* of a graph G, discussed by Dobrynin and Yuansheng et al., is the matrix $\tau(G) = \|\tau_{i,j}\|$, where $i = 1, 2, \ldots, p(G)$ and $j = 1, 2, \ldots, p(G) - 1$, and the value of $\tau_{i,j}$ is the number of paths with initial vertex v_i having length j. In this paper we seek to determine the minimal order p such that there exist nonisomorphic graphs G_1, G_2 with order p and $\tau(G_1) = \tau(G_2)$. The problem is posed for graphs in general, as well as for specialized families of graphs, including r-regular graphs, in which each vertex is connected to r other vertices. We construct a pair of nonisomorphic 5-regular graphs G_{20-1} and H_{20-1} of order 20 such that $\tau(G_{20-1}) = \tau(H_{20-1})$. We approach various aspects of the problem in several interesting manners, and furthermore attempt to generalize the results obtained for larger regularities in graphs.

Design and Characterization of a Littrow-Configuration External-Cavity Diode Laser

Wenxian Hong Under the direction of
Raffles JC Dr. Oskar J. Painter
Singapore Caltech

An enhanced tunable Littrow-configuration external-cavity diode laser (ECDL) is constructed from inexpensive commercial components and then characterized. The introduction of an additional mirror parallel to the diffraction grating allows the laser to be tuned without changing the direction of the output beam. The 785 nm ECDL has a typical power output of 4 mW at 56.0 mA of current input with an effective tuning range of 13 nm and exhibits excellent output-beam directional stability. An upper limit on its linewidth is placed at 63.0 MHz.

Improved Copper Resistance in *E. coli* Using a Heterologous Copper Chaperone from *Schizosaccharomyces pombe*

Andrew P. Klein Under the direction of
West HS A. James Link
Coralville, IA Caltech

In light of recent studies of cell surface labeling via the copper-catalyzed [3+2] reaction, it is desirable to engineer copper-resistant strains of *E. coli.* Whereas chromosomally encoded copper-detoxification systems focus on copper efflux from the cell, yeast possess proteins for copper detoxification that rely on copper sequestration. To this end, we have introduced the fourth domain of the *Schizosaccharomyces pombe* copper chaperone Pccs into *E. coli* via recombinant DNA techniques. XL-1 Blue cells expressing Pccs-IV possess improved copper tolerance as compared to the untransformed XL-1 Blue cells. We believe strains containing Pccs will be useful for further investigations into cell surface labeling via the [3+2] reaction.

The Effect of Chemical Additives on Ice-Spike Formation

Samuel Lederer Under the direction of
TJHSST Dr. Kenneth Libbrecht
Burke, VA Caltech

Solutions of sodium bicarbonate, calcium chloride, ethanol, acetic acid, and methanol, in concentrations ranging from 10^{-7} to 0.1 molar, were frozen under controlled conditions to measure the dependence of ice-spike formation on chemical impurities in water. The probability of ice-spike formation was plotted against molarity of solution to determine general trends for different solutes. The varying effects of different solutes at the same concentrations suggested that freezing-point depression was unlikely to cause inhibition of ice-spike formation in solutions, and a model based on dendritic sheet formation was proposed.

On a Generalization of the Collatz Conjecture

John Lesieutre Under the direction of
State College Area HS Mr. Zuoqin Wang
State College, PA MIT

In this paper we analyze a generalized version of the Collatz conjecture proposed by Zhang Zhongfu and Yang Shiming. We present a heuristic argument in favor of their conjecture and generalize a number of fundamental theorems from the original problem. We then obtain results related to properties of the generalized conjecture which do not arise in the original.

The Effect of Wind on the Emission of Grass Pollen

Grace Lu Under the direction of
The Walker School Dr. Richard Flagan
Kennesaw, GA Caltech

Despite common misconceptions, physical laws dictate that the pollen grains of wind-pollinated grasses should remain on the anthers after dehiscence until disturbed by an external force. Upon controlled observation of the flowering processes of rye-grass and Bermuda grass, it was observed that only 5–20% of the pollen escaped during dehiscence, leaving a majority of the pollen on the anthers. With the use of a miniature wind tunnel, it was discovered that a minimum threshold wind speed of 2.5 m/s was required to remove the remaining pollen. Experiments were performed with harvested flowers and artificial anthers made of pollen-coated substrates. The minimization of surface attractions resulting from nanostructures on the surfaces of the pollen and anthers, coupled with an absence of electrostatic forces, caused the velocity threshold for the natural and artificial anthers to be lower than that of the other surfaces dusted with pollen.

Glucocorticoid Regulation of Cell Survival in Normal and Cytokine-Treated Gastric Chief Cells

Jacob Sanders Under the direction of
AAST Dr. Kimihito Tashima
River Edge, NJ BIDMC

Gastric chief cells secrete pepsinogen and are found at the base of gastric glands. Chief cells constitutively express the antiapoptotic protein Bcl-$\mathrm{x_L}$, which may enable them to live longer than other gastric cells. We reveal that the cytokines released as part of the inflammatory response during *Helicobacter pylori* infection dose-dependently disrupt chief-cell monolayer integrity and viability. We also show that glucocorticoids dose-dependently increase expression of Bcl-$\mathrm{x_L}$ and hypothesize that this accounts for the ability of glucocorticoids to protect chief cells from cytokine-induced cell death. In addition, we show that glucocorticoids restrict the ability of cytokines to damage chief-cell monolayer integrity. Finally, we propose two mechanisms by which cytokines overwhelm the protective ability of glucocorticoids, resulting in chief-cell deletion and gastric atrophy during *Helicobacter pylori* infection.

Pointed Isoperimetric Inequalities

David Yang Under the direction of
Naperville North HS Mr. Huadong Pang
Naperville, IL MIT

The *Isoperimetric Theorem* states that for a planar region of given perimeter, the circle encloses the maximum area. In this paper, we determine the shape that maximizes area when the boundary is constrained to length L and forced to contain two points x units apart. For $L \geq x\pi$ maximum area is clearly attained by a circle of diameter $\frac{L}{\pi}$ containing the points. When $2x < L < x\pi$, we prove that the maximum area is attained when the boundary is two congruent circular arcs. We also consider the case of three equidistant points x units from each other. For some constant M, when $3x < L < M$, the maximum area is attained when the boundary is three congruent circular arcs. When $L \geq M$, the maximum area is attained when the boundary consists of two congruent circular arcs of length a and a third complementary circular arc of length $b \geq a$.

Abstracts of Other Reports

Terrain Analysis and Determination of Terrain-Specific Safe Speed for the Autonomous Vehicle Bob

Robin George Abraham Under the direction of
Delhi Public School Dr. Richard Murray
Kuwait, Kuwait Caltech

The focus of this research is to enhance the performance of Bob, the autonomous vehicle developed by the California Institute of Technology for participation in the DARPA Grand Challenge race. Bob was unable to complete the race in March 2004. Bob needs an effective system to analyze the terrain that it is traveling on and choose a safe speed for each that would allow it to successfully complete the race, yet at the same time will not cause damage to Bob. This paper examines the use of infrared sensors to analyze the terrain and also simulates a model based on which Bob could select a safe speed for each terrain, based on a safe stopping distance. This paper endeavors to make Bob a serious contender for the DARPA Grand Challenge to be held in October 2005.

A Worst-Case Delay Analysis for Opportunistic Transmission in Fading Broadcast Channels

Usman Ahmed Under the direction of
Salamat-ICAS Prof. Babak Hassibi
Lahore, Pakistan Caltech

Delay in transmission is one of the factors that demonstrate efficiency in a wireless network. We simulate the delay for the optimal throughput scheme and find its relation to the number of users. Our model is based on a single-antenna broadcast block-fading channel with n users with packet-based transmission. We showed that in a homogenous network, where the arrival rate equals the transmission rate, the maximum length of the queue increases without bound, and we illustrate the need for a device to bound the maximum length of the queue for large n.

Arabic Name Normalization via Machine Transliteration

Abdullah Al Mana Under the direction of
Najed HS Dr. David D. Palmer
Riyadh, Saudi Arabia Virage Adv Tech Group

The most important problem in transliterating names from Arabic into English is that there exists more than one way to write it in English. A name like “Qaddafi” has more than 49 possible English spellings. This causes a problem if we are searching for this name in English. We might not get any results if we typed it in using different English spelling than what is saved in the database. In this paper we present a method to connect all different English versions of an Arabic name via machine transliteration to match it with its Arabic spelling.

Caloric Restriction as a Treatment for Epilepsy in EL Mice

Stephanie Alessi Under the direction of
Mattanawcook Academy Dr. Thomas N. Seyfried
Lincoln, ME Boston College

Epilepsy is a neurological disorder whose causes are closely linked to irregularities in brain metabolism. Adjustments in diet can alter levels of glucose and ketones in the blood and regulate metabolism, thus preventing epileptic seizures. In this study, we examined the effects of caloric restriction on epileptic EL mice by administering a low-calorie diet for several weeks and testing weekly for changes in seizure susceptibility. Our observations are not yet complete, but the partial results we have acquired to this point support the hypothesis that a decrease in caloric intake lowers glucose levels and prevents seizures in EL mice.

Cloning a Bypass Suppressor of the Essential Histone Acetyltransferase Esa1

Priscilla Ang
Raffles JC
Singapore

Under the direction of
Dr. Michael F. Christman
Boston University

The histone acetyltransferase Esa1 is an essential yeast protein that has been shown to have many functions including transcription, regulation of cell-cycle progression, and DNA double-stranded break repair. The goal of this research is to determine the essential function of Esa1 using a bypass suppressor screen to identify a suppressor gene which can eliminate the need for *ESA1.* This suppressor gene has not yet been identified, though a possible candidate is *RAD57*, which is involved in meiotic DNA recombinase assembly.

Fe-Au Core-Shell Nanoparticles for Magnetic Cell Separation

Harold Au Shao Xian
Raffles JC
Singapore

Under the direction of
Dr. Mansoor Amiji
Northeastern University

Iron-gold core-shells are superparamagnetic nanoparticles which can be functionalized with biological molecules such as antibodies and fluorescent markers. Their size on the nanoscale allows them to be readily endocytosed. As such, these nanoparticles can be used to target specific cells for separation, for example tumor cells. By characterizing the magnetic nanoparticles in areas such as particle size, surface charge and UV spectrophotometry, suitability of Fe-Au core-shell nanoparticles for magnetic cell separation can be assessed. Red blood cell and tumor cells were labelled and separated successfully by an external magnetic field. The significance of this research is to demonstrate the viability of Fe-Au core-shell nanoparticles and pave the way for studies into tumor cell separation and specific, localized drug-delivery techniques as an alternative means to invasive chemotherapy.

Cascades of Responses to Disruption of Telomeres in Cultivated Human Melanocytes

Yael Ben-David
Amit Renanim
Ra'anana, Israel

Under the direction of
Dr. Hee-Young Park
Boston School of Medicine

UV irradiation causes a chain of reactions which protects the skin from further damage: formation of thymine dimers (pTpT) on the DNA complex and unravelling of the telomere tips (T-oligos), exposing a strand of DNA. This exposed strand of DNA induces NER (nucleotide excision repair), which frees the thymine dimers from the strand of DNA. In melanocytes, thymine dimers are free in the cellular space and p53 levels are reduced, which either causes cellular senescence or induces a differentiated phenotype which generates tyrosinase, a rate-limiting enzyme in the biosynthesis pathway of melanin. In order to discern the preliminary cause for the chain of reactions, the cells must be treated with a specific factor. UV irradiation causes many reactions, in addition to DNA damage. TRF2^{DN} specifically unravels telomeres and exposes 3' overhang. It has been shown by a student in Dr. Mark Eller's lab at the BU Medical Center that seven days after transfecting TRF2^{DN}, the level of melanin production increases. Thus, T-oligo mimics the effect of UV without the requirement of antecedent DNA damage. In this report, the cellular reactions to T-oligo and to TRF2^{DN} are compared, and the results show similar reactions induced by both.

An Interactive Game for Learning Robotics

Ahmad Berjaoui
Lycée Abdel Kader
Ghobeiry, Beirut

Under the direction of
Mr. Oren Zuckerman
MIT Media Lab

We present a system to teach basic robotics concepts to young children and to novices. With this system, users can interact with robots by modifying the robot's environment and the robot's characteristics. The systems' mechanics and the design of predefined and user-defined robots are discussed.

3D Imaging of Neurofibromatosis Tumor Growth Towards Treatment Enhancement

Shubha Bhat	Under the direction of
The Baldwin School	Dr. Gordon Harris
West Chester, PA	MGH

Bilateral vestibular schwannomas are benign tumors of the auditory nerves that are characteristic signs of neurofibromatosis type 2, a neurogenetic disease. Although these tumors are benign, they may put pressure on parts of the brain that control vital bodily functions. To understand the rate of growth of these lesions, 10 patient MRIs were evaluated both linearly and volumetrically, and these measurements were compared in the axial and coronal imaging planes. According to the results, the volumetric and linear analysis reveal similar trends, but the volumetric analysis represents the growth better and can be viewed in the coronal and axial views more accurately. In addition, the tumor in general grows faster outside the bone than inside the bone. Further studies examining tumor-volume growth in different parts of the head in a large patient pool may also contribute to better patient evaluations.

The Complexity of Summation

Ez-zobair Bidine	Under the direction of
Biranzarane	Mr. Vito Iacovino
Tahnaout, Morocco	MIT

In this paper, we study how the largest number of summations needed to calculate any set of sums made of n variables depends on the number of these sums. We start by determining the complexities $L(2,n)$, $L(3,n)$, and $L(m,n)$ of computing any m sums formed from n variables. We consider both real and $\mathbb{Z}/2$-valued variables.

The Use of Mass Spectrometry in the Detection of Oxidation Products in Alzheimer's Patients

Madison Capps	Under the direction of
Pineview Home School	Dr. Alfred Fonteh
Kennesaw, GA	Huntington Med Res Inst

It is suspected that oxidation of DNA plays a critical role in the development and progression of Alzheimer's disease. If this is the case, although the concentration of oxidation products in plasma, urine, and CSF would be elevated above that of a control subject without Alzheimer's, the actual concentration of oxidation byproducts would still be relatively low. The development of a sensitive and specific process to detect such low concentrations is an essential for all future research in this area. This study examines this problem and delineates a functional process to detect oxidation products using a DNA extraction method, solid phase extraction, and HPLC/ tandem triple quadrupole mass spectrometry.

Distribution and Detectability of Dark Matter in the Present Universe

Marissa Cevallos	Under the direction of
George Washington	Dr. Edmund Bertschinger
Charleston, WV	MIT

The existence of dark matter is one of the largest unanswered questions in cosmology. Dark matter has yet to be detected in a lab setting, even though it is theorized that ninety percent of matter in the universe is dark matter. Understanding how dark matter is distributed in the universe is crucial to detecting it; detecting dark matter is crucial to understanding the underlying physical laws of the universe.

Almost Monochromatic Triangles and Other Ramsey Problem Variants

Rohan Chabukswar	Under the direction of
Indian Inst Tech Bombay	Mr. Aaron Tievsky
Pune, India	MIT

In this paper, we investigate variants of the Ramsey-number problem. In particular we study chromatic graphs not containing almost monochromatic triangles and finite metric spaces not containing almost equilateral triangles. We study the relations between the different answers to these problems. We study the upper bounds and lower bounds for these numbers and how they are related to each other and the original Ramsey numbers.

Synthesis of 3-amino-2-alkenoate

Albert Chen
Lexington HS
Lexington, MA

Under the direction of
Dr. Mohammad Movassaghi
MIT

The main aim of the project is to synthesize an isomerically pure sample of 3-amino-2-alkenoate, a useful β-amino acid derivative that is a precursor to molecules with medicinal properties. We successfully made the product from the reaction of ethyl acetoacetate and ammonium acetate dissolved in methanol. We also attempted to synthesize 3-pyrrole-2-alkenoate from the reaction of our product and 2,5-dimethoxytetrahydrofuran. The products of both reactions were chracterized by means of ^{1}H nuclear magnetic resonance spectroscopy.

Development of a Connectivity Algorithm to Assess the Quality of Tissue-Engineered Bone

Jason Chen
OK School of Science and Math
Edmond, OK

Under the direction of
Dr. Brian Snyder
BIDMC

A novel approach to healing bone fractures has been developed: seeding mesenchymal stem cells (MSCs) onto porous silk scaffolds and providing fluid-flow stimulation with spinner flasks. This tissue-engineered (TE) bone must be mechanically tested for strength. Before mechanical testing, however, the connectivity of the bone must be assessed to determine whether the actual strength of the bone matrix, and not the scaffold, is being tested. We have developed an algorithm to evaluate this property. Results of idealized model testing show that it is a substantial improvement over existing analytical methods. This algorithm has future applications in determining the strength of TE bone and predicting fractures.

The Least Number of 1's

Sisi Chen
Northview HS
Duluth, GA

Under the direction of
Mr. Mohsen Bahramgiri
MIT

Through addition and multiplication, the number 1 can be used to express any integer. We let $f(n) =$ the least number of ones necessary to express the number n. As $n \to \infty$, also $f(n) \to \infty$, but $f(n)$ is not always increasing. Sometimes $f(n+1) < f(n)$, and moreover sometimes $f(n) \leq f(n+1) - 2$, indicating that subtraction indeed helps us use fewer number of ones to express n. Let $g(n)$ represent the least number of ones required to express the number n using addition, multiplication, and/or subtraction. Clearly $g(n) \leq f(n)$. In this paper, we will prove some upper and lower bounds on both $f(n)$ and $g(n)$ and also provide evidence that there exists an infinite number of n for which $g(n) = f(n)$ and an infinite number of n for which $g(n) < f(n)$.

Overpayment in Strategyproof Payment Schemes

Yifei Chen
W Windsor Plainsboro N
Plainsboro, NJ

Under the direction of
Ms. Evdokia Nikolova
MIT

We consider the random graph model $G(n,p)$, a probability space of all graphs $G = (V, E)$ such that $|V| = n$ and each edge appears with probability p. The random graphs we consider are weighted, with each edge having unit cost. We give a computational method that determines the expected Vickrey-Clarke-Groves (VCG) overpayment $Q_E(n,p) = \langle \frac{T_P - T_C}{T_C} \rangle$ explicitly for a random biconnected graph on n nodes. We use this method to generate overpayment functions for $3 \leq n < 8$. For larger n, we use random simulations of graphs to obtain a plot of $p \mapsto Q_E(n,p)$ in $G(n,p)$. We also consider power-law random graphs in which the number of nodes with degree k is proportional to $k^{-\beta}$ for some $\beta > 2$. We study VCG overpayment empirically in power-law random graphs similar in structure to the Internet using the preferential-attachment model, a graph process that adds vertices and edges to an existing graph. This model generates a power-law random graph with $\beta = 3$.

Impact of Impaired Mitochondrial DNA Repair in the Brain

Allen Cheng
Arcadia HS
Arcadia, CA

Under the direction of
David Simon, M.D., Ph.D.
BIDMC

Somatic mitochondrial DNA (mtDNA) point mutations are hypothesized to play a role in age-related neurological dysfunction. Somatic mutations can occur by oxidative damage, either directly to the mtDNA or to free nucleotides. However, several repair mechanisms operate within the mitochondria to prevent against mutation, such as those initiated by *OGG1* and *MTH1*-coded enzymes. Using the cerebral cortices of *OGG1* knockout (KO), *MTH1* KO, and double KO mice, we screened for mtDNA mutations and compared the mutational burden to that of wild-type mice. We also assessed dopaminergic neuronal density in each genotype by tyrosine-hydroxylase (TH) staining. Although staining results were inconclusive, we found a six-fold increase in mutational burden in double KO mice relative to wild-type mice. Our results suggest a large impact of DNA repair deficiency on mutational burden, thus providing a model for possible protective measures against the accumulation of mtDNA mutation.

Model for Anomalous Properties of Water

Debbie CHIANG Under the direction of
Garden City HS Prof. Udayan MOHANTY
Garden City, NY Boston College

In this work, we generalize to three dimensions a statistical thermodynamic model of water based on the Truskett and Dill model. Our model stipulates that water molecules can exist in three distinct states: expanded, dense, and cagelike. The isothermal-isobaric partition function of a collection of N water molecules is determined, enabling us to calculate the probability of a triplet, or “cell,” of water molecules being in a certain state relative to the temperature and pressure of the system. We predict the molar density, thermal expansion coefficient, and isothermal compressibility as functions of temperature and pressure. The predictions are shown to properly reflect available experimental data.

Some Results on Inclusive and Exclusive Partitions of Complete Graphs

Robert CORDWELL Under the direction of
Manzano HS Dr. Richard WILSON
Albuquerque, NM Caltech

The partitioning of graphs is a well-known and difficult but useful problem. Results in this field often can be used to solve otherwise difficult combinatorial problems. Here we consider the complete graph K_n, where the vertices are evenly spaced around a circle C. We show that, with such a representation, any K_n for n odd can be partitioned into a set S_{inc} of noncrossing closed paths such that the intersection of their convex hulls contains the center of C. If n is odd and sufficiently large, we show that K_n can be partitioned into the set of noncrossing closed paths S_{exc} such that the union of their convex hulls does not include the center of C. We also examine some combinatorial results which relate to these problems.

Automatic Classification of Music Through Acoustic and Lyrical Analysis

Ruth DHANARAJ Under the direction of
Carmel HS Dr. Beth LOGAN
Carmel, IN Hewlett-Packard Labs

In this paper, we explore the use of digital analysis of music in order to identify likely hit songs. We use both acoustics and lyrics alone and a mixture of the two to extract features from the music in an attempt to identify these songs. Preliminary results seem to show that lyrics are more useful than the audio stream in predicting hit songs.

A Constant Dilation Field of the Low-Energy Effective Action Implies Compactness of the Underlying Manifold

Connemara DORAN Under the direction of
Doran Home School Ms. Fangyun YANG
Potomac, MD MIT

In this paper, we examine a question inspired by a section of Perelman’s recent proof of Thurston’s famous Geometrization Conjecture. Perelman demonstrates that if a manifold is compact (historically called closed), then the minimizer (dilation field) of the low-energy effective action functional is constant. We conjecture that the compactness of the manifold is also *necessary* for the dilation field to be constant. For all two-dimensional noncompact surfaces of revolution in $\mathbb{R}^3$, we prove that all minimizers must be nonconstant, and prove that we can construct such minimizers.

Investigation of Gait Change in Larval *Drosophila*

Omer Durak Under the direction of
TEV Inanc Turkes HS Dr. Daniel Tracey
Gebze, Turkey Caltech

The larval *Drosophila* is an extensively studied organism and is an appropriate animal to examine in order to find answers to questions about the brain. In this study, wild-type and mutant *Drosophila* larvae were used to examine movement after stimulation by noxious heat. Worm Tracker was used for recordings and stimulation analyses. Results showed that wild-type larvae move with turns, stops, and side-to-side head probes before stimulation. After stimulation, the larval paths were straight and the linear speed increased. Unlike wild-type larvae, *ppk1* mutants moved with decreased number of turns and stops before and after stimulation. The results presented here provide a basis for future movement studies of *Drosophila* larva.

Evaluation of Vented Fuel for an Advanced Nuclear Reactor

Hussein El-Bolbol Under the direction of
Rafic Hariri HS Prof. Michael Driscoll
Saida, Lebanon MIT

The feasibility of venting a nuclear fuel assembly to prevent excessive buildup of internal pressure without releasing a significant amount of radioactivity is examined. A time-dependent diffusion equation which incorporates radioactive decay is solved. Important findings are that a vent tube 400 cm long prevents the escape of all significant radioactive species with the exception of ^{85}Kr and 133Xe. CO_2 coolant is significantly better than He, and high pressure is beneficial.

Thiolated Gelatin Nanoparticles: A Novel Approach for Intracellular Delivery in Gene Therapy and Drug Delivery

Joline Fan Under the direction of
Upper Arlington HS Ms. Sushma Kommaredy
Columbus, OH Northeastern University

In this study, thiolated gelatin nanoparticles were tested against unmodified gelatin nanoparticles for robustness and precision as vectors for drug delivery and gene therapy applications. The thiolated gelatin nanoparticles have a positive surface charge, which assists in targeting negatively charged tumor cells. Trafficking of nanoparticles in NIH 3T3 cells was studied using a fluorescent dye as the payload. The thiolated nanoparticles were observed to release the dye instantly upon entering the cell due to the reduction of their disulfide bonds, thereby yielding a uniform distribution of dye throughout the cell. The gelatin nanoparticles, on the other hand, remained intact and slowly released their payload through degradation. Similar results were observed with a plasmid DNA payload, as measured by green fluorescent protein expression from the plasmid. These results show that thiolated gelatin nanoparticles are more effective than gelatin nanoparticles for intracellular delivery of drugs and DNA, which may thus lead to more promising cancer therapies.

Coverage Analysis for Circular Periodic Orbits

Nickolas Fortino Under the direction of
Phillips Academy Dr. Martin W. Lo
North Andover, MA Jet Propulsion Laboratory

The coverage-analysis problem for nonperiodic circular orbits can easily be solved by numerically calculating a double integral. For periodic orbits, no such simplification exists, and the orbit must be projected over one period to calculate the exact answer. However, the same double integral can be used to give an approximate answer, and evaluating the double integral is about three orders of magnitude faster than obtaining the exact answer. With an appropriate measure, the maximum error from this method can be expressed in terms of this measure, independent of the minimum elevation angle of the ground station and the planet the satellite is orbiting.

Structure Learning for Inductive Inference

Eugene Frenkel Under the direction of
St. John's School Dr. Joshua Tenenbaum
Houston, TX MIT

Our study has a dual focus in what we model and test: how do people organize the information they know about a certain knowledge domain into a structure, and what inductive inferences can they make on the basis of such a structure? We hypothesize that people have a general capacity to do both for any domain. We present a formal model on that basis and compare it to two preexisting models that do not consider any explicit domain knowledge such as structure. We developed and implemented a methodology for experimentally testing structure learning and property induction in an abstract domain where only a domain-general structure-learning ability would apply.

On a Conjecture of Chowla and Chowla

Andreas GALANIS
Psychico College
Athens, Greece

Under the direction of
Dr. Kimball MARTIN
Caltech

In this paper, we treat some special cases of a conjecture of Chowla and Chowla, finding quadratic polynomial conditions on a prime so that the period of the continued fraction expansion of its square root is of a given length. We also find the numbers p for which $\sqrt{p} = [x, \overline{v, v, \ldots, v, 2x}]$ holds. Finally, we estimate the period length of special quadratic numbers.

Synthesis of the Germanium(IV) Corrole

Jason GONSALVES
Manchester HS
Manchester, CT

Under the direction of
Dr. Karn SORASAENEE
Caltech

The recent discovery of an efficient synthetic method for corroles without requiring a solvent or a catalyst has created new opportunities to study the unique properties of metallocorroles. However, a pure sample of the chloro-coordinated germanium(IV) corrole (Ge-Cl-(tpfc)) has never been obtained. This paper describes the method by which a germanium(IV) corrole may be synthesized.

Thrust Optimization in Underwater Synthetic Jets

Maxwell GRAZIER G'SELL
Saline HS
Saline, MI

Under the direction of
Dr. Joel BURDICK
Caltech

The thrust output from synthetic jets, possibly a more effective alternative to propellers for precise maneuverability of underwater vehicles, is examined. A system for optimizing parameters of the jet's driving function using a genetic algorithm is established. Optimal parameters of a sawtooth driving wave, including the frequency, amplitude, and tooth angle, were examined. Improvements to the process are discussed, along with possible future applications.

Regulation of Late-stage Flower Development by Downstream Genes of the Homeotic Protein AGAMOUS

Jerry GUO
Riverside HS
Greer, SC

Under the direction of
Dr. Toshiro ITO
Caltech

The genes and pathways AG regulates in late-stage flower development are largely unknown. Several putative downstream genes of AG involved in anther dehiscence were identified, including DAD1, MYB26, OPR3, COI1, and RBR1, through a bioinformatics approach. A 35S::AG-GR inducible line was constructed for timed-induction expression analysis of anther dehiscence versus indehiscence, which revealed the likely direct induction of DAD1 by AG. Mutated AG binding sites of DAD1 and MYB26 were linked to the β-glucuronidase reporter gene; the plasmid constructs were transformed into plants for GUS staining to test *in vivo* site activity. Results support the hypothesis that AG is continually functional and controls late-stage flower development by regulation of downstream genes.

Investigation of the Changes in the Composition of Dendritic Spines When N-methyl-D-aspartate (NMDA) Receptors are Stimulated

Jiaqi Guo	Under the direction of
Raffles JC	Prof. Mary Kennedy
Singapore	Caltech

Dendritic spines are specialized postsynaptic structures where most excitatory synapses in the brain are found. They play an indispensable role in information storage, as past research demonstrates that any alteration to spine morphology corresponds to changes in the strength of synaptic connections, which are proposed to be the essential substrates of memory. This project aimed to identify some of the modulators of spine morphology that are regulated in an activity-dependent manner, the activity in this case being the stimulation of N-methyl-D-aspartate (NMDA) receptors, which has been observed to cause the expansion or collapse of dendritic spines depending on the duration and intensity of stimulation. In particular, this project investigated whether the concentrations of synGAP and Kalirin and the activity of cofilin in the postsynaptic density change in response to a brief stimulation of NMDA receptors and hence determine how NMDA regulates spine morphology. It sought to begin the definition of the molecular pathway that links synaptic activation to changes in spine morphology, the basis for the treatment of neurological disorders. Murine hippocampal tissue was treated with NMDA, and purified using a postsynaptic density preparation. Then, a BCA assay for protein determination, followed by Western blotting, was done before protein level analysis by densitometry was carried out. The results obtained confirmed that NMDA treatment causes certain changes in the protein composition of the postsynaptic density and consequently spine morphology.

Characterization of Structure-Function Homology in Recently Isolated Piscine Zα Protein Domains

Kelley Harris	Under the direction of
C.K. McClatchy HS	Dr. Ky Lowenhaupt
Sacramento, CA	MIT

Zα is a winged helix-turn-helix protein domain that binds specifically to Z-DNA, a left-handed isomer of the classic double helix. Functionally conserved Zα domains exist in organisms as diverse as poxviruses and vertebrates, although the biological role of Z-DNA remains unknown. Here, we assay wild-type and recombinant Zα domains isolated from novel piscine PKZ proteins for Z-DNA-binding activity. The results of our mutagenesis study imply that the Zαs newly isolated from piscine DLMs differ significantly from previously studied Zαs in their conformations when bound to Z-DNA and in the residues they use to bind the helix directly. This finding implies that we must look to the tertiary structures of Zα wings to elucidate the true evolutionary history of this clade. An understanding of the sequence drift, functional conservation, and convergent functional adaption that created the extant families of Zα domains could give us valuable insight into universal mechanisms of protein evolution.

Temperature and Magnesium-Dependent Pause-Point Analysis in DNA Constant-Force Unzipping

Sarah Harrison	Under the direction of
Charlottesville HS	Dr. Mara Prentiss
Charlottesville, VA	Harvard

Unzipping double-stranded DNA (dsDNA) into single-stranded DNA (ssDNA) yields information that can be used to better understand and predict DNA replication and binding. dsDNA under the constant applied force of magnetic tweezers unzips in a series of pauses and jumps caused by energy barriers which occur when the energy of the bonded dsDNA cannot be overcome. The pause points of DNA constant-force unzipping were tested under varying temperatures in a phosphate buffer solution (PBS) as well as in a magnesium solution. Sequence-dependence was evident in the durability of the pause points at all temperatures, even under strong forces. Lambda phage DNA was unzipped in both directions to confirm evidence of sequence-dependence. The magnitudes of the jumps in the magnesium solution were slightly shorter than those in PBS because the DNA condensed in the presence of magnesium, but the data suggested that the pause points are located in the same positions along the DNA molecule as in PBS, which is consistent with the theory that magnesium shortens DNA. The temperature-varying experiments showed that the temperature at which dsDNA is stored has an effect on the energy of the nucleotide bonds.

Synthesis of a Model Oblate Colloidal Solution

Steven Hershman Under the direction of
Lower Merion HS Dr. Rebecca Christianson
Wynnewood, PA Harvard

We explored three techniques used to create particles for use in a model oblate colloidal system. Model colloidal systems have a number of implications for both industry and our fundamental understanding of how atoms and molecules behave. By compressing polystyrene spheres dried on brass sheets at high temperatures, we were unable to form oblates without causing the particles to fuse together. Red blood cells were inflated using osmotically controlled solutions and stiffened by crosslinks, but were not uniform. We found that by compressing hot poly(vinyl alcohol) films with embedded sulfated polystyrene particles, we were able to produce reasonably uniform oblate particles for use in model colloidal systems.

Calcium and Magnesium Independence in R-cognin-Mediated Retina-Cell Adhesion

Ryan Jamiolkowski Under the direction of
Cary Academy Dr. Robert E. Hausman
Cary, NC Boston University

Retina cognin (R-cognin) is a cell surface protein that mediates the tissue-specific adhesion of embryonic chick retina cells. R-cognin has disulfide isomerase activity that is known to be required for the promotion of cell adhesion, but its substrate is not known. The identification of R-cognin's substrate(s) and the characterization of its adhesion pathway are important tasks in unraveling the mechanisms of retinal differentiation and development. Protein interaction cloning has narrowed the potential substrates to myocilin, members of the integrin family, and members of the tetraspanin family. To determine whether the target is an integrin or any other adhesion molecule dependent on the presence of divalent cations, suspended retina cells were placed in a medium lacking calcium and magnesium to observe whether the substrate of reaggregation was changed under those circumstances. We have shown here that retina cell adhesion at embryonic day 8 (E8), which depends in part on the R-cognin-mediated pathway, occurs even in the absence of Ca^{2+} and Mg^{2+}. This result suggests that the adhesion pathway does not include the divalent-cation–dependent integrins.

Automated Face Recognition: The Application of Artificial Intelligence to Computer-Vision Systems

Camden Jansen Under the direction of
Pacific Coast HS Dr. John Fisher
Santa Ana, CA MIT AI Lab

In the paper, we discuss the two major difficulties faced by current face-recognition systems. We then present a solution that converts pictures into eigenfaces and combines affine transformations with a search function to analyze these transformed images. We quantify the increases in search speed and the improvements in match results.

GSTT1 Deletions are Associated with the Progression of Meningiomas in an Adult Neurofibromatosis Type 2 Population

Wei Ying Jen Under the direction of
National JC Dr. Mia MacCollin
Singapore MGH

Neurofibromatosis 2 is an autosomal dominant disease which predisposes people towards developing intracranial tumors. Meningiomas are a common form of NF2-related brain tumor which are hard to cure due to their inaccessible location on the brain. Here, we investigate a potential link between the gene *GSTT1*, involved in the metabolism of environmental toxins, and the initiation and progression of meningiomas. We have found that although *GSTT1* is not associated with meningioma initiation, it is involved in meningioma progression. Our findings provide a basis for further investigation of the genetic basis of NF2-related tumors and potential mechanisms for control of tumor growth.

Semiconductor Nanowire Growth

Lewis Jones Under the direction of
Arnold School Dr. Jagadeesh Moodera
Lytham, UK Francis Bitter Magnet Lab

In this paper, the effects of different bases on semiconductor nanowire growth are studied. A tube furnace was used to grow the nanowires by vapor-liquid-solid deposition. It was found that silicon bases were better for creating regular nanowire growth, especially for the indium oxide nanowires. Gold colloid was tested for its substrate effectiveness and was found to be highly effective.

An Investigation of the Effects of a Quark-Matter Core on Neutron-Star Cooling

Pooja Jotwani Under the direction of
Charles W. Flanagan HS Prof. Krishna Rajagopal
Pembroke Pines, FL MIT

Observation of cooling neutron stars can potentially provide information about the states of matter at supernuclear densities. Physical properties important for cooling, such as neutrino emission processes and specific heat quantities inside and outside the star, are reviewed. The neutrino processes include the modified Urca process and the direct Urca process for nucleons and exotic states of matter, such as quark matter. The dependence of theoretical cooling curves on the assumed internal composition of the star and observations of thermal radiation from isolated neutron stars are described. Quark-matter cores of varying sizes are used to predict the effect of exotic matter on the cooling of neutron stars.

An RNAi Screen for Meiotic Genes in *Caenorhabditis elegans*

Stefan Kathman Under the direction of
Summit Country Day Dr. Monica Colaiacovo
Covington, KY Harvard Medical School

Meiosis is the process by which organisms that reproduce sexually divide their chromosomes to produce haploid gametes. Microarray data was analyzed to determine 192 genes most likely to be involved in meiosis. We then used RNA interference (RNAi) to selectively knock down three of these genes. The RNAi screen is being used to narrow down the pool of genes most likely to play a critical role in the synaptonemal complex, which forms in prophase I of meiosis. We observed the RNAi-treated worms and their progeny for phenotypes (such as a high embryonic lethality or a high incidence of males) indicating a meiotic defect. We also fixed and stained the worms for cytological analysis to determine whether there were any defects in germline morphology or chromosome number. One of the three genes we tested, Y45F10D.9, produced sterile worms and abnormal gonad morphology, suggesting that it is potentially involved in meiosis.

Assessment of Deep Boreholes for Disposal of Nuclear Reaction Spent Fuel

Hazem Khayat Under the direction of
Rafic Hariri HS Prof. Michael Driscoll
Saida, Lebanon MIT

Recent studies of deep boreholes have indicated that this approach is a very promising way to get rid of the highly radioactive wastes produced by nuclear power plants. This research deals with evaluating the features including the rock type and the minimum depth for deep boreholes to ensure complete isolation of these wastes from the surface. In the process of this study, it was found that igneous rocks such as granite or crystalline with depth between 500 and 2000 m can trap the most radioactive waste materials for a very long period of time: considerably beyond 10^6 years, and thus preventing any radionuclides from rising up to the surface and causing harm to the environment and living things.

Squares Touching a Constant Number of Other Squares

Mohammad Khderat Under the direction of
Jubilee School Mr. Zuoqin Wang
Zarqa, Jordan MIT

We consider collections of squares with the property that each square touches exactly n others along part of an edge. We also generalize this idea to other shapes, such as circles and rectangles, and to more inclusive definitions of "touching." We have explored one of the open problems in this field, proving that there is no collection of 6 squares each of which touches exactly 3 other squares. Also, we have

generalized this theorem and used it to formulate other results involving different shapes.

Role of *fgf8* and *fgf24* in Craniofacial Development of Zebrafish

Tyler Kieft Under the direction of
Essex HS Dr. Pamela Yelick
Essex Junction, VT Forsyth Institute

Fibroblast growth factor (fgf) signalling proteins regulate growth, proliferation, and specialization of cells in developing vertebrate embryos. Here we investigate the role that *fgf8* and *fgf24* play in craniofacial development of zebrafish, *Danio rerio.* We perform geometric morphometric analyses to determine the statistical significance of mutated phenotypes. Zebrafish that are homozygous for the mutated form of *fgf24*, called *ika*, show moderate differences in phenotype, while fish that are homozygous for the mutated form of *fgf8*, called *ace*, show major defects in craniofacial structure. Double mutants either have severe deformation of craniofacial structures or fail to form any craniofacial cartilage. We also hypothesize that heterozygotes have slight craniofacial differences from the wild-type, but the experimental sample size is too small to determine whether these differences are statistically significant.

Carbon Nanotubes

Koganti Lahari Under the direction of
Mahindra UWC Prof. Mildred S. Dresselhaus
Vijayawada, India MIT

The properties of one-dimensional carbon nanotubes (CNT's), which are similar to a two-dimensional graphene layer, have attracted much attention from researchers. Raman spectroscopy is used to compare Raman shifts (frequencies) between cut and uncut SWNT's, allowing one to study the electronic structure and properties of the SWNT's. This Raman shift is caused by the creation or absorption of phonon modes including the RBM, D-band, and G-band. The only variable between uncut and cut semiconductor SWNTs is length. We have observed no effect of length on semiconductor SWNT's when the laser light is cross-polarised. Not changing the positions of the cutting lines which are wavevectors allows only certain wavelengths of energy to be in resonance with the transition energy of the valence and conduction bands.

Near-Infrared Albedo and Flux Variation of Saturn's Rings at 2.00 μm: 1995–2003

Johann Komander Under the direction of
AR Math & Sci Dr. P. A. Yanamandra-Fisher
Greenwood, AR Jet Propulsion Laboratory

Voyager and ground-based observations have established water ice as the prime constituent of Saturn's rings. The reflectivity of light at 2.00 μm corresponds to a narrow water-ice absorption band and provides valuable information regarding particle size, distribution, and ring thickness. Albedo, a measure of the reflectivity of solar incidence, is used to study chemical absorption patterns of water and indicates increased particle size and water-ice proportions as radial distance from Saturn increases through comparisons to known chemical absorption patterns. Furthermore, comparisons of ground-based albedo data from 1995, 1997, and 2003 suggest that maximum reflectivity is located between ring opening angles of $-0.6°$ and $-26°$ from the equatorial plane. The albedo variations in each of the Rings A, B, C, and Cassini Division during the observational period show that the brightness of the rings increased from 1995 to 1997, decreased from 1997 to 2003, and globally fluctuates at the east and west ansae. These results are consistent with the reflectivity of rotationally locked nonspherical ring particles or purer ice in the rings' interior. Distinct differences in the rings' albedos support the Cassini-Huygens data showing purer ice moving away from the planet.

Universality Properties of Positive-Definite Integral Quadratic Forms

Scott Kominers Under the direction of
Walt Whitman HS Mr. Aaron Tievsky
Bethesda, MD MIT

In this paper, we present several results in the representation theory of quadratic forms. We prove the uniqueness of the 2-universality criterion provided by Kim, Kim, and Oh and find a criterion for 3-universality of positive-definite integral rank-6 quadratic forms. We define and examine the properties of lattices that represent rank-n forms in multiple ways. Finally, we generate a specific set of rank-2 forms and develop a criterion for universality for the set.

Collagenolytic Properties of Enamelysin (Matrix Metalloproteinase 20)

Justin Komisarof
The Shipley School
Haverford, PA

Under the direction of
Dr. John Bartlett
Forsyth Institute

Matrix metalloproteinase 20 (MMP 20), also known as enamelysin, is a unique MMP found only in the tooth enamel organ and that is a likely factor in tumor metastasis in oral cancers. We investigated its substrate specificity by use of a peptide library, which suggested that enamelysin would cleave type V and type III collagen, two previously unknown substrates. Enamelysin was found to cleave collagen type V, and probably type III as well. We also found that enamelysin does not cleave type I collagen in its native form, but it does cleave it in the denatured form. These new findings about enamelysin's substrate specificity help to reveal its possible ability to enable metastasis in oral cancer.

Brake Analysis and Design of an Autonomous Vehicle

Michael Laures
Naperville Central HS
Naperville, IL

Under the direction of
Dr. Richard Murray
Caltech

Caltech's autonomous vehicle, Bob, has an inadequate braking system. In order to complete and win the 2005 DARPA Grand Challenge race, it must have improved braking capabilities. The main delays with the current system are caused by computer delay and a slow brake actuator.

Purification of the *Drosophila* Telomere-Associated Protein Gayani

Leow Shi Ling
Hwa Chong JC
Singapore

Under the direction of
Dr. Svetlana Rashkova
MIT

Telomere maintenance has been implicated in diverse fields from aging to tumorigenesis. The telomere maintenance mechanism in *Drosophila melanogaster* relies on the retrotransposons Het-A and TART. It is suspected that regulatory proteins interact with them to control the rate of telomere elongation. The novel protein Gayani has been identified as a candidate regulatory protein that associates with telomeric regions and enables the successful telomeric targeting of retrotransposons. Here we present a method for purifying the Gayani protein and discuss its optimization.

Determining the Relationship Between Local and Global Spatial Processing in Vision

Randolph Li
Montgomery Blair HS
North Potomac, MD

Under the direction of
Dr. Jason J.S. Barton
BIDMC

The brain uses a combination of componential and noncomponential (holistic) processes when analyzing visual stimuli. The purpose of this investigation is to determine which system is utilized more by the brain and to determine the effects of rotation on both systems. The experiments present stimuli with either destroyed local or global frames of reference in a search for an alteration in the stimulus. We expected to find that local frames of reference were more useful than global frames of reference and that rotation negatively affected spatial ability in both local and global frames of reference. The results, however, suggest that componential and holistic processing are interchangeable in the visual system, and that rotation only affects global frames of reference.

Evaluation of *cis-cis* HOONO Production by Laser Photolysis Through Modeling of Reaction Kinetics

Da Lin
Westview HS
Beaverton, OR

Under the direction of
Dr. Mitchio Okumura
Caltech

Detection and study of *cis-cis* HOONO under atmospherically relevant conditions are expected to significantly influence current atmospheric models. Recent proposals have been made to utilize laser photolysis to produce *cis-cis* HOONO for spectroscopic detection under atmospheric pressures. A model was created using Facsimile to evaluate the feasibility and determine the initial conditions necessary for experimental detection of *cis-cis* HOONO at

30, 100, and 500 torr of atmospheric pressure and 298 K. Studies identified 16 reactions that are influential in NO_x activity in an experimental reaction cell. Simulations based on ordinary differential equations show that the percent *cis-cis* HOONO yield should be 5.27%, 7.57%, and 14.1% respectively, which are in good agreement with past theoretical and experimental data. Analysis of predicted models identifies the importance of several free-radical concentrations for experimental studies.

Closure Properties of D_{2p} in Finite Groups

Po-Ling Loh Under the direction of
James Madison HS Dr. Michael Aschbacher
Madison, WI Caltech

We discuss the generalized version of a question posed by the topologist E. Farjoun about closed embeddings of a finite group H in a finite group G. We study the case $H \cong D_{2p}$ for p an odd prime and determine a sufficient condition for $H = G$ when H is closed in G.

Rational-Distance Sets on $y = x^2$

Dawn Mackey Under the direction of
Half Hollow Hills HS West Ms. Fangyun Yang
Wheatley Heights, NY MIT

A collection of points in space is a *rational-distance set* if the distance between each pair of points is rational. One may ask if it is possible to find rational-distance sets of N points on the parabola $y = x^2$. In this paper, we characterize all the sets of five points at rational distance on $y = x^2$. Extending the results of Garikai Campbell, we use elliptic surfaces to prove the existence of infinitely many rational-distance sets of five points on $y = x^2$. Next we discuss what conditions would restrict four points at rational distance to the right half of the parabola. Afterwards, we prove that because rational-distance sets exist on $y = x^2$, so do integral-distance sets. Lastly, we examine rational-distance sets on other smooth curves, such as the hyperbola $y = 1/x$ and the family of parabolas in general. In particular, we prove that there are infinitely many rational-distance sets of four points on every parabola with rational coefficients.

Characterizing Martian Soil Analogues with Optical Microscopy

Paul Magyar Under the direction of
Classical HS Dr. Kimberly Kuhlman
Providence, RI Jet Propulsion Laboratory

Characterization of soil is an important part of any exploration of the Martian surface. By studying the shapes of small soil particles, it is possible to learn about the large-scale geologic processes to which they have been exposed. Two-dimensional image analysis was here used to characterize sand particles by their origin. Shape parameters were used to compare the shapes of particles. This information was then used to identify the geologic type of another sample and to characterize a mixture of sand types. The eventual objective of these studies is pseudoconfocal microscopy, through which an optical microscope can create three-dimensional images of particles and characterize their shapes.

The Effect of Vestures on Membrane Deflection in the Xylem Pits of Vascular Plants

Alan Malek Under the direction of
The Harker School Prof. N. Michele Holbrook
Los Gatos, CA Harvard

The ability of plant xylem vessels to carry water without allowing gas to infiltrate them depends on the properties of the microscopic pit connections between the vessels. It has been suggested that vestures, which are finger-like projections from the cell wall of a pit, would limit the ability of the porous membrane covering these pits to stretch under pressure by acting as a physical barrier. We used measurements of the air-seeding pressure (the pressure needed to move a gas bubble across the pit membrane) to calculate the membrane pore size, which in turn allowed us to estimate the membrane stretch in different species with and without vestures. We found that membrane stretching was not reduced by the presence of vestures.

An Automated Training Apparatus for the Study of Memory in Planaria

Olga Mandelshtam	Under the direction of
University HS	Dr. Michael Levin
Irvine, CA	Forsyth Institute

Planaria are an ideal model for the study of memory and learning in higher organisms. To date, experiments with planarian training have been impractical and inefficient. We have designed the Automated Training Apparatus (ATA) to remedy the flaws of most past planarian memory studies. We have set up the hardware and the software to capture images of the planaria, detect the worms, and output their positions. This functionality is crucial to the planarian memory project for which the ATA was designed.

News-Broadcast Boundary Identification Using Feature Density

Advay Mengle	Under the direction of
Newport HS	Dr. David Palmer
Bellevue, WA	Virage Adv Tech Group

News broadcasts contain a vast amount of information. In order to find relevant sections of video, search engines must provide users with precise news results, such as specific news stories related to a user's query. This requires that boundaries between stories and commericals be identified in the broadcast and the content between boundaries be categorized. This paper presents a novel trainable method of identifying the boundaries using the density of certain features in the broadcast. The implementation of this method turned out to be a reliable identifier of real boundaries, but at the same time incorrectly marked extra boundaries.

An *Ab Initio* Study of Amino Acids in Vacuum and Aqueous Solution

Caleb Ng	Under the direction of
Libertyville HS	Prof. Nicola Marzari
Libertyville, IL	MIT

Amino acids are vital to life in their role as the monomers that form proteins. Many amino acids have recently been found to have special uses in medical treatments as well. Therefore, elucidation of their most stable structures is often very useful, since structure and function are very closely tied in biological systems. Using *ab-initio* methods, powerful computers were used to model such molecules by solving, or approximating solutions to, complex quantum-mechanical equations that would otherwise be unsolvable. In addition, a molecular-dynamics simulation was performed to demonstrate the changes in the properties of the amino acid glycine in aqueous solution, as it is often found in biological systems.

The Roles of Bone Morphogenetic Protein-3 in the Regulation of Cartilage Formation

Nghi Nguyen	Under the direction of
Bishop O'Dowd HS	Dr. Laura Gamer
Oakland, CA	Forsyth Institute

During embryo development, the geometry of skeletal elements is laid down by the formation of cartilage (chondrogenesis) as a template. Two crucial proteins controlling early limb chondrogenesis are bone morphogenetic proteins (BMPs) and activins. In this study we have examined the expression of the signaling components used by BMPs and activins during limb chondrogenesis and found that the BMP-3 signaling pathway and its components are a synexpression group, a group of genes that function in a common pathway and are expressed in the same tight temporal and spatial pattern. We also found that BMP-3 influences both the BMP and activin signaling pathways. BMP-3 does so by behaving as a pseudoligand, interfering with the pathway by competitively binding to the ActR-IIA/B receptor, which is used by both BMPs and activin. These data suggest that BMP-3 can effectively block activin signaling by binding to its only type II receptor, ActR-II, and eliminating activin signal transduction. In contrast, BMP-3 may actually amplify BMP signaling by blocking BMP ligands from binding to ActRII and making them available to bind BMPR-II. BMP-3 thus appears to have a dual role in regulating chondrogenesis in early limb-bud development.

N-glycosylation as a Regulator of Cell Proliferation

Eva Nong Under the direction of
Miami Palmetto Senior High Dr. Aneta Liwosz
Pinecrest, FL Boston University

E-cadherin, a transmembrane glycoprotein found in many intercellular junctions, functions as a tumor suppressor, which prompts further research. Because E-cadherin needs to be modified by N-glycosylation to function properly, mammalian tumor cells often exhibit a weakened version of E-cadherin with aberrant glycosylation. HTB cells were treated with five different inhibitors, extracted in β-octylglucosidase buffer, and examined using SDS-PAGE and Western blot. Protein lysates were analyzed on film, showing lower N-glycosylation in cells after both wortmannin and FTI-277 inhibitor treatment. The results suggest that the role of N-glycans in tumor proliferation can be regulated by downstream signaling pathways. We hope our research will ultimately lead to some preventative measures in cancer treatment.

Thermal System Optimization for the Mars Gravity Biosatellite

Mark A. Norsworthy Under the direction of
Lakenheath American HS Col. John E. Keesee
RAF Lakenheath, England MIT

The Mars Gravity Biosatellite Project will launch 15 mice into space and observe them for 35 days at Martian gravity levels. One of the challenges of the mission is the maintenance of a stable internal temperature for the welfare of the mice and integrity of the payload systems. The systems inside the payload module, in addition to the mice themselves, generate heat which must be transferred to the satellite bus for radiation into space. A test thermal assembly was constructed, and preliminary results show that the design will meet requirements. The requisite internal temperature was 22.3°C and the desired temperature of the top of the aluminum rod was 15°C.

Field Emission Study of Single- and Double-Walled Carbon Nanotubes

Christopher Palmer Under the direction of
Kubasaki HS Dr. Zhifeng Ren
Okinawa, Japan Boston College

Chemical vapor deposition was used to grow carbon nanotubes from powder catalysts. Samples grown using the different catalysts were analyzed using SEM, then tested for field emission. The catalyst using a Fe:Mo:MgO ratio of 5:5:60 was able to reach a practical tunneling voltage with the smallest electric field.

Knights on a Chessboard

Richard Pang Under the direction of
Baton Rouge Magnet HS Mr. Huadong Pang
Baton Rouge, LA MIT

We evaluate the minimum number of moves it takes for a knight to travel from one point on a board to another. We consider a variety of board shapes and sizes, including some that are not planar. In doing so, we can generate a formula that applies for many different dimensions and also provide a bound for the minimum number of moves it takes to move to any given point.

Spontaneous Activity in Spiking Neural Networks with Global Inhibition and Sparse Lateral Excitation

Eric Paniagua Under the direction of
Massapequa HS Dr. Dezhe Jin
Massapequa, NY MIT

We extend the work of Jin and Seung to networks of N excitatory neurons subject to global inhibition. We describe the dynamics of the network using continuous- and discrete-time Markov chains. We find the probabilities of specific sequences of states occurring. We find the entropy of the sets of k-length state sequences generated by networks.

The Effects of Atmospheric Turbulence on Simulated Starlight in a Low-Pressure Telescope Environment

Renee Park Under the direction of
Montgomery Blair HS Dr. Benjamin Lane
Spencerville, MD MIT

We investigated the feasibility of using nulling-interferometry techniques in the stratosphere to search for extrasolar planets. Studies have shown wavefront distortions from atmospheric turbulence to be too large to obtain clear ground-based images. Wavefront errors caused by atmospheric turbulence in the stratosphere are on the scale of 1 angstrom; however, the level of the wavefront error inside a telescope tube is unknown. We used Kite, a system level testbed created by the Jet Propulsion Laboratory for the External Metrology System of the Space Interferometry Mission, to simulate atmospheric-turbulence conditions inside a telescope tube in the stratosphere. We obtained 40 minutes of laser-metrology gauge readings in relatively stabilized temperatures at 5 mb. The data was analyzed using adaptive optics and blanking techniques to reduce the level of instrument noise. The integrated total error was approximately 1.849 angstroms, validating the possibility of applying nulling-interferometry techniques without sending a telescope into space.

Dopamine, Serotonin, and Brain-Derived Neurotrophic Factor Receptors on Spindle Cells in the Frontoinsular Cortex of Humans

Corinne Pender Under the direction of
TJHSST Dr. John Allman
Annandale, VA Caltech

Past research has demonstrated that neurotransmitters dopamine and serotonin are associated with the reward system of the brain and implicated in aggression, respectively. Studies have also shown that brain-derived neurotrophic factor (BDNF) is necessary for memory formation and neural plasticity. This study investigates the presence of these receptors on spindle cells in the frontoinsular cortex. The finding that these receptors are present on the spindle cells in the area demonstrates that spindle cells may play a role in the reward system, aggression, and memory formation and may help to explain autism and other disorders that are characterized by deficiencies relating to these functions. This experiment also investigates the difference between spindle cells and pyramidal cells, another type of neuron. These two types of cells were found to be quantitatively different in terms of feret and y-skew.

Normal Coefficient of Restitution in Oblique Collisions of Two Macroscopic Particles in a Viscous Fluid

Sam Powell Under the direction of
Sidwell Friends School Dr. Melany Hunt
Chevy Chase, MD Caltech

A system of two pendulums was used to perform oblique macroscopic particle collisions in a viscous fluid. The normal coefficients of restitution of these collisions were tested as a function of the impact Stokes number. We investigate whether the normal coefficient of restitution in an oblique collision can be predicted as a function of the Stokes number in the normal axis, independent of the collision angle. Data from the experiment did not indicate a global difference in the collisional dynamics of the normal component of an oblique collision and a normal collision.

The Formation of the Spin of Mars

Eric Price Under the direction of
TJHSST Dr. David J. Stevenson
Falls Church, VA Caltech

Current theories for the formation of terrestrial planets have trouble explaining the presence of Mars. The inability to produce objects similar to Mars through N-body simulations has led to the hypothesis that Mars is a leftover planetary embryo and has suffered no giant collisions. The spin of Mars is used as a guide to evaluate various formation scenarios. The spin of Mars is numerically simulated during formation, assuming a power-law mass spectrum and isotropically oriented velocities in the planetesimal swarm. It is found that a Mars-sized planetary embryo is likely to have approximately the same angular momentum as Mars, while collisions between embryos tend to cause more spin than that of Mars.

On the Solvability of $\mathfrak{p}$-adic Diagonal Equations

Antoni Rangachev Under the direction of
Paisii Hilendarski HS Mr. Mohsen Bahramgiri
Sofia, Bulgaria MIT

This paper deals with some fundamental questions in the study of the diagonal diophantine equation $a_1x_1^k + \cdots + a_sx_s^k = 0$ over a finite extension K of the field $\mathbb{Q}_p$ of p-adic numbers, namely some new upper bounds on the number of variables needed to ensure their solvability.

Social Signals: Interfaces Designed to Improve Group Interaction

Christopher Reed Under the direction of
Quantico HS Ms. Joan DiMicco
Quantico, VA MIT

When one person dominates group discussions, the other members are unable to give their own opinions. This leads to a lack of communication as well as a lack of knowledge. When each member of the group is able to contribute ideas, the group is able to share more knowledge and is capable of making better decisions. The purpose of this experiment is to design an interface that will promote more equal participation in meetings in order to make groups more capable of making informed decisions.

Common-Sense Knowledge Extraction: A Web Search Algorithm

Kimberly Reinhold Under the direction of
St. Joseph Jr.-Sr. HS Dr. Push Singh
Hilo, HI MIT Media Lab

A computer program designed to extract common-sense knowledge from the world wide web was written in the language Python. Using a three-step algorithm (search web for query string, extract related verb phrases, rank output), this program could find many verb phrases correlated with input queries. Three main ranking methods were investigated and compared: statistical, common-sense, and a combination of statistical and common-sense. Output could be fed into an XML-RPC web script from which users could rate concept pair relations. Information gathered from the user was stored with the data extracted from the web. The knowledge obtained could then be used in a variety of tasks. The project applies common-sense computing in a novel way and contributes to other research in the MediaLab, including work on StoryNet and ConceptNet.

A More Accurate Approach to Closed-Handed Computer Bridge

Alex Schwendner Under the direction of
Home School Dr. Christopher Umans
Austin, TX Caltech

The field of computer bridge has been somewhat underdeveloped. Game-tree search techniques can be and have been applied to the game of bridge, but current applications are lacking in that they assume, for the sake of speed, perfect information for all plays after the current play. In this paper, we present a new technique which better accounts for imperfect information and which, although slower, we argue is still feasible and should yield considerably better play.

Evolution and the Out-of-Africa Theory: Novel Analysis Using a Multipopulation Data Set

Nicholas P. Semenkovich Under the direction of
Ladue HS Dr. David Reich
St. Louis, MO Harvard Medical School

The most prevalent theory of human evolution is the out-of-Africa theory that humans evolved from a common African ancestor. It describes a population within Africa which expanded to the rest of the world. The theory is well documented regarding expansion out of Africa; however, little is known regarding structuring within the ancestral African populations. While the out-of-Africa theory assumes that there was one homogenous population that expanded outside of Africa, these data are not well supported. This project describes a novel computational analysis of a multipopulation SNP data set in order to better understand the out-of-Africa expansion.

Dependence of Collective Plasmon Resonance Frequency on Silver Nanoparticle Array Geometry

David Sheets
Paul Laurence Dunbar HS
Lexington, KY

Under the direction of
Dr. Harry Atwater
Caltech

Electron plasma oscillation (plasmon) resonance frequencies in arrays of silver nanoparticles are affected by a variety of factors. Among these are interparticle spacing and array length. Thus, the resonance can be tuned throughout the technologically interesting visible and infrared spectra by manipulating these properties. Currently, fabrication of these arrays is infeasible using conventional lithographic methods. To better understand the effect of polarization, interparticle spacing, and chain length on the dominant resonance frequency in a chain of 5-nm-radius particles, electromagnetic simulations were done. These simulations show that in nanoparticle arrays of length greater than or equal to 4 the resonant frequency is not affected by addition of more nanoparticles when driven by transversely polarized light.

The Effectiveness of the STEREO Mission Using Current Imaging Technology

Parth Sheth
Memphis Univ School
Germantown, TN

Under the direction of
Dr. Paulett Liewer
Jet Propulsion Laboratory

NASA's upcoming Solar Terrestrial Relations Observatory (STEREO) mission consists of two solar-imaging spacecraft which will provide stereo images of the Sun. One use of these stereo images will be to create 3D models of the Sun. This paper analyzes the accuracy of current 3D solar-modeling programs to determine the effectiveness of the 3D-modeling aspect of the STEREO mission and finds that accurate 3D models can be generated for a wide range of spacecraft separation angles, with the best results at spacecraft separation angles around 45°.

Learning by Playing: Creating Custom-Board-Games Framework

Alexander Simeonov
Math HS P. Beron
Varna, Bulgaria

Under the direction of
Prof. Walter Bender
MIT

The purpose of this project is to create a basic framework that will allow players to create their own board games by designing the boards and the pieces, defining the rules and the winning conditions. Further on, the games will be tested against other human players, and when suitable winning strategies are discovered and formalized, the users will be able to design their own computer opponents. So far, a simple two-players game of checkers has been implemented.

A Plunge into Supercooled Liquids: Relaxation and Inhomogeneity

Paweł Śledź
High School No. 13
Szczecin, Poland

Under the direction of
Prof. Udayan Mohanty
Boston College

In this paper, the Xia and Wolynes model of random first-order description of supercooled state is generalized. We show an intimate connection between this model and the classic Adam-Gibbs configurational entropic model. Fluctuations of the number of molecules within a cooperative rearranging region are included in the Adam-Gibbs model for the first time. The fragility index is computed and compared with available experimental data for polymeric glass formers.

Photochemical Decomposition of Pyruvic Acid in Ice

Miranda Stewart
University HS
Tucson, AZ

Under the direction of
Dr. A. J. Colussi
Caltech

The photochemical decomposition of pyruvic acid, $CH_3C(O)C(O)OH$ (PA), in aqueous solutions and ice produces CO_2. Acetoin, $CH_3C(O)CH(OH)CH_3$, is also formed, but in significantly smaller amounts than expected. We thoroughly searched for unidentified reaction products in D_2O media using gas-chromatographic analysis with mass-spectrometric detection. A mechanism is proposed for the title reaction.

Polynomial-Time Approximation Algorithm for a Traveling Salesman Problem with Binary Costs

David Stiebel Under the direction of
Jewish Acad. of Metro Detroit Mr. Joungkeun Lim
Bloomfield Hills, MI MIT

There is no known polynomial-time algorithm to solve the traveling-salesman problem (TSP). Approximation algorithms exist for certain cases of the TSP. For a TSP with binary costs, the greedy algorithm finds a tour with cost no worse than 1.75 times the optimal cost of a tour. An improvement of the greedy algorithm presented in this paper finds a tour with cost no worse than 1.66 times the optimal cost of a tour. The generous algorithm is presented, and a probabilistic argument is made claiming that it finds a tour with a cost significantly nearer the optimal cost. Data gathered from comparing the generous algorithm with the improved greedy algorithm supports this claim.

Investigations in Cake-Cutting

Zuoyu Tao Under the direction of
Stuyvesant HS Mr. Joungkeun Lim
New York, NY MIT

This paper is about cake-cutting, the problem of fairly dividing a cake among n-player when all n players have different valuations of the cake. We present a few new algorithms for three and four players utilizing money and multiple cakes. We also review the current progress on envy-free cake-cutting procedures. Finally, we give some results on exact cake-cutting and highlight its importance to envy-free cake-cutting.

On Families of Sets with Restricted Intersection Values

Dmitry Taubinsky Under the direction of
Maria Carrillo HS Mr. Alan Leung
Santa Rosa, CA MIT

Let $\mathcal{A} = \{A_1, A_2, \ldots, A_m\}$ and $\mathcal{B} = \{B_1, B_2, \ldots, B_m\}$ be two collections of subsets of an n-element set. Given certain intersection properties that must hold between the collection $\mathcal{A}$ and $\mathcal{B}$, in 2003 Snevily conjectured a bound on m. We prove his conjecture for two special cases, prove a bound on m for another case, and provide useful techniques for working with the general case. We will use probabilistic methods to find lower bounds on m for similar systems of sets.

A Study of Very Soft X-Ray Sources in the Milky Way

Shriharsh Tendulkar Under the direction of
Indian Inst Tech Bombay Dr. Rosanne DiStefano
Mumbai, India Harvard-Smithsonian Center

We describe the search for Galactic quasisoft sources based on a 188 Ks Chandra ACIS-I observations of a patch of the Milky Way in the constellation of Scutum. The observation was spilt into three parts of 37 Ks, 59 Ks and 97 Ks. We detect 8, 7, and 15 sources satisfying the spectral criteria for quasisoft sources in the three observations respectively. However, they are much lower in luminosity than quasisoft sources detected in external galaxies.

Searching for High-Redshift Candidates from the Epoch of Reionization

Daniel Thai Under the direction of
North Canyon HS Dr. Jeremy Darling
Phoenix, AZ Carnegie Institute of Washington

Recent data taken from the NRAO VLA Sky Survey (NVSS) and the Westerbork in the Southern Hemisphere (WISH) survey have revealed ultra-steep-spectrum radio sources. These spectrums suggest high redshifts. An analysis of ultra-steep-spectrum K-band 2.2-micron images was conducted to find estimated redshifts of these objects. Two sources that are likely from the epoch of reionization were detected. This analysis of candidate objects is part of a plan to probe the epoch of reionization.

Preparation of *B. subtilis* Membranes To Induce Bacteriophage ϕ29 DNA Ejection in Solution

Zachary Travis Under the direction of
North Shore HS Dr. Rob Phillips
Glen Head, NY Caltech

Understanding the methods by which bacteriophages inject their DNA into host cells and the role of pressure and enzymatic mechanisms in viral DNA transport is vital to understanding the infection process of viruses. Currently, direct observation of phage ejection in solution can only be accomplished with select types of phages, such as λ and T5. To accurately model and understand the method by which viruses eject their DNA, information from more viruses is needed. For example, models of packing and ejection forces for ϕ29 bacteriophages have predicted pressures and ejection characteristics, but since the virus currently cannot be made to eject in solution, actual testing of these predictions is impossible. In this study the ϕ29 host cell, *B. subtilis*, is lysed so that the specific receptor protein for the virus is kept in solution. This preparation leaves the receptor proteins in their original conformation, allowing them to bind to ϕ29 in solution and stimulate DNA ejection. This research enables direct observation of DNA ejection by ϕ29 and can be applied to other *B. subtilis* and Gram-positive bacteria.

Application and Consequences of Surface-Energy-Based Cloth Simulation Methods

Ryan Tsukamoto Under the direction of
Iolani School Dr. Peter Schröder
Honolulu, HI Caltech

Creating realistic computer-generated images with virtual cloth is a challenge for both artists and graphics programmers. Cloth is impossible to model using the physics of rigid bodies. Though it is difficult, including plausible virtual cloth in computer-generated scenes adds a sense of realism. We propose a simple system for describing the geometry and physics of cloth that can be fine-tuned for use in a variety of situations.

Statistical Approach to Background Noise in Gamma-Ray-Burst (GRB) All-Sky Spectrometer Experiment (GASE)

Kaya Tutuncuoglu Under the direction of
TED Ankara College Dr. Miguel F. Morales
Ankara, Turkey MIT

This paper aims to determine statistical ways to decrease errors on radio telescope experiments. It focuses on background noise and radio frequency interference, which are two main sources of deceiving signals. For background noise, a method of averaging multiple data is considered; and for RFI, the importance of excision is considered.

Passdoodle: a Lightweight Authentication Method

Christopher Varenhorst Under the direction of
Wichita Collegiate School Mr. Max van Kleek
Wichita, KS MIT

This paper investigates the use of unique finger traces, or *doodles*, as a means of authentication in a pervasive environment. Velocity here is investigated as means to uniquely identify a doodle. A blurred distribution grid created from combined training samples and the variance across this grid are also used for recognition. These three systems used together have produced accurate results for a population of ten users. The research presented here may have applications to handwriting and drawing recognition as well.

Saccadic Inhibition Deficits in Schizophrenia: An MEG-Based Study

Veena Venkatachalam Under the direction of
Governor Livingston HS Dr. Dara S. Manoach
Berkeley Heights, NJ MGH/Harvard

Past studies have demonstrated that patients with schizophrenia have trouble performing the antisaccade task, which requires subjects to inhibit the automatic prosaccade response. The role of the frontal eye fields (FEF) in this inhibitory mechanism has also been studied, and the correlation between FEF activation and inhibition has been documented. The current study presented both healthy subjects and those with schizophrenia with a saccadic paradigm requiring inhibition and compared the temporal patterns of FEF activation (determined by magnetoencephalography, or MEG) of the two groups. The results indicated that the right FEF was activated about 500 ms before the antisaccade was made in a representative subject with schizophrenia and about 100ms before the antisaccade was made in a representative healthy subject. In addition, the FEF was activated to a greater extent in the patient afflicted with schizophrenia. These observations—that the subject with schizophrenia showed earlier and more intense FEF activation when performing the same task as the healthy subject—led to the conclusions that persons with schizophrenia have saccadic inhibition deficits that cause them to work longer and harder than their normal counterparts to perform the same antisaccade and that this difficulty is manifested in greater FEF activation because the FEFs play a key role in saccadic inhibition. This study lacked the statistical power necessary to draw definitive conclusions; however, if future studies do verify these preliminary results, the antisaccade task could become an instrumental tool in everything from researching the pattern of schizophrenia inheritance to identifying those at risk of developing schizophrenia.

Room-Temperature Ferromagnetism in Germanium-Manganese Thin-Film Semiconductors

Lynwood Walker Under the direction of
Hahnville HS Dr. Jagadeesh Moodera
Hahnville, LA Francis Bitter Magnet Lab

Ferromagnetic semiconductors are semiconductors in which some of their constituent elements are replaced by a magnetic impurity such as Mn. Semiconductors that are ferromagnetic at room temperature could be used to inject spin into circuits so as to function along with charge in electronic operations. GeMn is predicted to be a ferromagnetic semiconductor at room temperature. We study the growth and the properties of germanium-manganese thin films. Films are grown by e-gun and thermal evaporation of the compositions 60/40, 85/15, and 97/3 germanium to manganese. Measurements of magnetization as a function of temperature and magnetic field were taken, and it was revealed that all three compositions of GeMn were ferromagnetic at room temperature. Measurements of resistivity versus temperature also showed that all materials exhibited semiconducting behavior. We discovered that GeMn could potentially be used as a ferromagnetic semiconductor in electronic applications.

Genetic Regulation of Forebrain/Midbrain Volume in Mice

William Austin Webb Under the direction of
Counterpoint Academy Dr. Glenn D. Rosen
Custer City, OK BIDMC

Hypothesizing that forebrain/midbrain (F/M) volume in mice is a heritable trait that can be linked to a number of genes, we used stereology to estimate the F/M volumes of 142 BXD recombinant inbred mice. We then used WebQTL to perform a marker regression (visualized on an interval map) against F/M volume and locate probable quantitative trait loci (QTLs), pairwise scanning to find epistatic interaction among loci, and various bioinformatic databases to identify potential genetic determinants of F/M volume. Suggestive QTLs were located on chromosomes 7, 18, and 19, but 7 and 19 were later found to be brain-weight determinants/QTLs and thus were eliminated. Analysis of chromosome 18 yielded five genes: four members of a family of protocadherin-encoding genes (the *Pcdha* family) involved in cell/cell adhesion and neural network formation and one *wnt*8 gene that is expressed in the neural plate in early development.

Generation of an Ubchain-Thiirane Metalloisopeptidase Inhibitor

Jingyuan Wu Under the direction of
Oak Ridge HS Dr. Rati Verma
Oak Ridge, TN Caltech

Enzyme complexes called 26S proteasomes and the complementary targeting molecules called ubiquitins comprise eukaryotic cells' major ATP-dependent system for degrading and destroying endogenous proteins. RPN11, a subunit of the 26S proteasome, is believed to be responsible for a crucial, deubiquitinizing step in this proteolytic process. This project focused on the generation of a thiirane-based inhibitor that is to be used as a stronger test of RPN11's functional significance. The results revealed several difficulties in utilizing an intein-based mechanism for generating the proteasomal targeting domain of the thiirane-based inhibitor, including possibly a novel and complicating reaction mechanism of the self-splicing intein domain. Eventual application of research done on this inhibitor may include the synthesis of cancer-treatment drugs that function by interfering with proteasomal activities.

Merger Rate of Galaxies at $z \lesssim 6$

Christopher Wulff	Under the direction of
Gym. Bad Aibling	Dr. Christopher Conselice
Bad Aibling, Germany	Caltech

We use the ACS UDF images to identify B-drop, V-drop and i-drop Lyman-break galaxies. We use the CAS system to analyze these galaxies and simulate high-redshift galaxies to calculate corrections for the CAS parameters. The merger fraction is then calculated and analyzed for $z \approx 4$, $z \approx 5$, and $z \approx 6$. The calculated star-formation rate is compared to the asymmetry of the galaxies and indicates that a higher star-formation rate in galaxies undergoing mergers may be the case.

A Common-Sense Approach to Object Recognition

Yuetian (Peak) Xu	Under the direction of
Stuyvesant HS	Mr. Ian Eslick
Woodside, NY	MIT Media Lab

In this paper, we discuss the application of a common-sense system to scene recognition. Traditional approaches for object recognition and scene recognition are driven by bottom-up information without any contextual information. Use of common-sense knowledge to provide top-down constraints in conjunction with traditional approaches will result in better recognition. Here a proposal for making a bidirectional system of recognition is made. In addition, a novel contour-matching process is discussed that allows for robust soft-matching of scene contours to a database of object contours. This makes contours more effective in scene recognition.

Can Average Color Guide Visual Search? A Study of Statistical Cueing

Diana Ye	Under the direction of
Timberline	Dr. Jeremy Wolfe
Boise, ID	Brigham and Women's/Harvard Med

Visual search is something people perform constantly, from searching for their car in the parking lot to finding a friend's face in the crowd. A big question in vision research is what guides visual search. Our study investigated whether the statistical properties of a scene could guide visual search, an effect termed statistical cueing. We compared studies in which the statistical property of color predicted for target location with studies of contextual cueing, in which the spatial layout of items in the target's context predicted for target location. The results indicate that statistical cueing using color does not have an effect with these stimuli. In the course of this work, we replicated the contextual-cueing effect of Chun and Jiang, but interestingly, our contextual-cueing experiments did not exhibit a slope effect between set sizes. This suggests that context might be facilitating not visual search, but only the decision-making process.

An Evaluation of the General Electric Medical System Functool Software Package

Owen Yeh	Under the direction of
Skyline HS	Dr. Gordon Harris
Sandy, UT	MGH

MIP imaging is an effective scanning modality for pinpointing cerebral aneurysms. The purpose of this investigation was to evaluate the effectiveness of the new General Electric Medical System Functool software package, a semiautomated maximum-intensity-projection (MIP) imaging software program. The quality of the new MIP software images was compared to that of MIPs created manually by 3D lab technicians. To do so, we measured the specificity, sensitivity, and speed of a neuroradiologist searching for aneurysms using both the new MIP software images and the traditional 3D lab images. The results suggest that the MIP imaging software is effective in all but the PCOM and cavernous sinus regions. These areas are easily analyzed using standard CT source images, so the MIP software images could simply be combined with the source images. Furthermore, because of its extremely fast creation time, the MIP imaging software provides an effective initial review prior to full MIP reconstructions.

Creating BGF Reader for the Computational Materials Design Facility

Yue Ying
Xin Zhong Senior HS
Shanghai, China

Under the direction of
Dr. Peter Meulbroek
Caltech

The Computational Materials Design Facility, a computational simulation environment, is now being developed in the Materials and Process Simulation Center at Caltech. The main difficulty is the gap between different software packages for sharing data. This paper develops a program called `BGF reader` which enables data sharing between software packages.

The mRNA Distribution and Functional Analysis of Zebrafish Caveolin-3

Bruna Zacka
Coll. Notre-Dame de Jamhour
Metn, Lebanon

Under the direction of
Dr. Pamela Yelick
Forsyth Institute

The three genes of caveolin are supposed to play an important role in the development of the organism. This research focuses on the function of caveolin-3 in zebrafish. To investigate the zebrafish caveolin-3 mRNA expression pattern, whole mount in situ hybridization was performed; then to deterime its role in embryonic development, caveolin-3 antisense morpholino oligomers were utilized to knock down caveolin-3 protein expression. We have found that caveolin-3 mRNA was strongly expressed in the somites and notochord. Targeted depletion of caveolin-3 by antisense morpholino oligomer showed systemic defects, including but not restricted to the brain, eye, ear, tail, and circulatory system in the morphants; larger dosage is lethal. To the contrary, the embryos injected with 4-mismatch caveolin-3 morpholino developed normally. Hence we have found that caveolin-3 plays indispensable roles in embryonic development and organogenesis.

Design of a Lighting System for the Mars Gravity Biosatellite

Yu Zheng
ECNU Senior High No. 1
Shanghai, China

Under the direction of
Col. John E. Keesee
MIT

The mission of the Mars Gravity Biosatellite is to build a biosatellite and send a small population of mice into low Earth orbit.

The living environment for the mice should be similar to natural conditions, including the air supply, water supply, food supply and lighting of the living space. This paper offers a design of the lighting system for the biosatellite.

A simple electrical circuit changes the intensity and the wavelength of light every 12 hours. Several types of light-emitting diodes (LEDs) have been tested and the most suitable one was picked. A simple prototype of the lighting system has been constructed by the author.