THE UCLA Undergraduate Science Journal

VOLUME 35 | SPRING 2022

THE UCLA UNDERGRADUATE SCIENCE JOURNAL

THE UCLA Undergraduate Science Journal

VOLUME 35 | SPRING 2022

Established 1986.
The UCLA Undergraduate Science Journal is an officially recognized student organization at UCLA.
The UCLA Undergraduate Science Journal is entirely managed and produced by students.
Its contributions are the original work of UCLA undergraduates.

FRONT COVER ILLUSTRATION BY CHELSEA LAI
INSIDE COVER IMAGE COURTESY OF ESO/Y. BELETSKY

A Letter from the Editors

Dear Reader,

We are incredibly proud to present to you the 35th Edition of the UCLA Undergraduate Science Journal. Over the past year, our student body has continued to endure the unpredictability of the COVID-19 pandemic while also experiencing dramatic growth with our return to campus in the fall. We have also been able to witness the power of scientific innovation, reflected not only in discoveries used to combat the pandemic, but also in the tenacity of undergraduate students who have had to adjust to a new normal while still putting out phenomenal research. We are honored to be a part of this scientific collaboration and we are confident you will see these values exemplified in this year's volume of the USJ. Whether as an author or a staff member, the entirety of this journal is the direct result of undergraduate students, from the articles themselves to the peer review, editing, and layout of the journal. We are truly impressed with the caliber of research conducted by the undergraduate authors, as well as the tireless diligence and dedication of the staff members in our Review, Editorial, and Layout Boards.

This volume of the USJ includes both original research and review articles, ranging from research on the metabolic effects of long-term prostate cancer treatment to the quantum mechanics of strong field electron emission. The articles featured in this journal may only offer a small glimpse into the breadth of research that undergraduates partake in at UCLA, but the range of topics covered is a true representation of the ingenuity and diversity of our students.

This year has been a memorable one for all of us. Many of us came back to Westwood for the first time in more than a year, and all of us had our first in-person classes for the first time since early 2020—an exciting but rocky process of readjustment. Our staff continued to operate remotely in light of the ever-changing public health conditions. Despite these challenges, we are extremely inspired by the passion of the student body in its commitment to scientific progress. The 2022 Undergraduate Research Week held virtually in May was a resounding success, featuring more than 1,300 total live and prerecorded student presentations, 110 live panels, and 24 multimedia exhibits.

This journal would not be possible without the labor and love of our entire team. We would first like to thank Colornet Press for helping us produce this beautiful journal. We sincerely thank Alec Hernandez, Program Coordinator of the URC–Sciences, and the UCLA Clinical and Translational Science Institute for their generous continued support. We would also like to especially commend the passion and dedication of all of our USJ staff members who continued to work hard to assemble the best journal possible during a time marked by uncertainty and change. A huge thank you goes to our managing editors: Kevin Jiang of Review, Sophia Yu of Editorial, and Nicole Lin of Layout, and the assistant managing editors for the amazing work and long hours you have all dedicated to the USJ. We are endlessly grateful to Dr. Janet Goins, our former faculty advisor for their guidance this year and for the past several years—we wish you the best in St. Louis! Last but certainly not least, we would like to express our most heartfelt thanks to our fearless faculty advisor and director of URC–Sciences, Dr. Tama Hasson, without whom the USJ would not exist.

We hope you are inspired while reading and learning about the science and research featured in this journal, in fields that touch many aspects of our daily lives. This journal stands as a testament to the sheer courage and determination of our UCLA journal staff and undergraduate authors—we could not be more honored to share it with you.

Warm regards,
James Yoon & Victoria Li

VOLUME 35 | SPRING 2022

TABLE OF CONTENTS

Mr. Frosty the Gecko Gives Insight into Skin Cancer

Aishwarya Atmakuri

Mr. Frosty the lemon frost gecko gained prominence when he was purchased for $10,000 by a breeder, and subsequently had offspring with several tiny white tumors. 80% of lemon frost geckos develop a skin cancer from iridophores, which are pigment producing cells found in reptiles. A team of researchers at UCLA found that the semi-dominant, spontaneous mutation gene SPINT1 is responsible for not only the sunny yellow color of the geckos, but also their skin cancer. In fact, SPINT1 has been shown to be implicated in skin cutaneous melanoma and over-proliferation of skin cells in humans. The effects of losing SPINT1 is dose-dependent and has led to tumors in other species as well. Tumors caused by losing SPINT1 could lead to new models of skin cancer, which will aid in understanding the SPINT1 pathway, examining differences in healthy geckos, and identifying potential treatments.

Image: Lemon Frost Gecko
Reference: L. Guo, J. Bloom, S. Sykes, E. Huang, Z. Kashif, E. Pham, K. Ho, A. Alcaraz, X. G. Xiao, S. Duarte-Vogel, L. Kruglyak, Genetics of white color and iridophoroma in "Lemon Frost" leopard geckos. PLOS Genetics. **17**, e1009580 (2021).
Image Source: Fauzan Maududdin – stock.adobe.com

Landmark in Fusion Energy Achieved – Burning Plasma

Nathaniel Chen

"Fusion energy is only 30 years away" has been a running gag for at least the last half century. However, this may not be the case anymore after the recent "burning plasma" achievement at Lawrence Livermore National Laboratory's (LLNL) Nuclear Ignition Facility (NIF). Burning plasma is when energy produced by fusion is high enough to self-propagate the reaction. To accomplish this, 1.9 megajoules of x-rays were directed by 192 lasers into a small hohlraum containing a perfectly spherical capsule filled with deuterium-tritium (DT) fuel. The symmetric bath of x-rays then reacted with the capsule to implode upon itself via ablation. This implosion pressure was enough to ignite the DT fuel, producing 1.3 megajoules of fusion energy, or about an 800% increase in energy compared to past inertial fusion trials. This puts fusion at the brink of net energy ($Q \geq 1$), which could revolutionize clean, safe, and compact energy production.

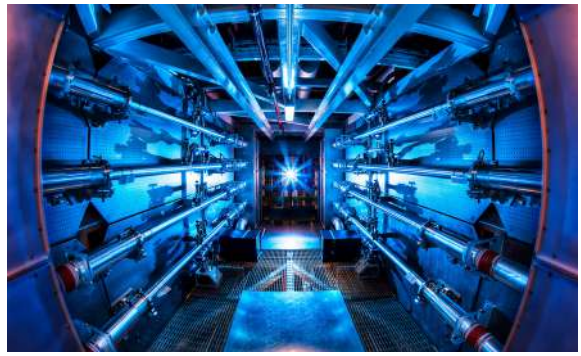


Image: Nuclear Ignition Facility (NIF) preamplifier chamber with a peak power of 500 terawatts.
Reference: Zylstra, A.B., Hurricane, O.A., Callahan, D.A. et al. Burning plasma achieved in inertial fusion. Nature 601, 542–548 (2022). https://doi.org/10.1038/s41586-021-04281-w
Image Source: "Preamplifier at the National Ignition Facility" by Damien Jemison/LLNL. Licensed under CC BY-SA 3.0. https://commons.wikimedia.org/wiki/File:Preamplifier_at_the_National_Ignition_Facility.jpg

A Dozing Diagnosis of Alzheimer's

Lauren Markus

Researchers at Tel Aviv University discovered a potential indicator of Alzheimer's that can be detected as early as twenty years prior to typical observable symptoms. By studying mouse models of Alzheimer's, researchers observed unusually high activity in the memory region of the brain, even while the mice were asleep or under anesthesia. This finding indicated that the normal homeostatic mechanisms responsible for maintaining a set point of neuronal activity had failed, similar to 'quiet' seizures in epileptic patients. Accordingly, the researchers treated the mice with a known epilepsy drug, which successfully suppressed this atypical brain activity and could prove beneficial as a future treatment for Alzheimer's. The researchers also intend to incorporate EEG monitoring into standard surgical procedures to analyze patient brain activity under anesthesia and determine if the same identification factors used in mice can predictively diagnose Alzheimer's.

Image: Photo portraying degeneration of memory in Alzheimer's patients.
Reference: Reference: D. Zarhin, R. Atsmon, A. Ruggiero, H. Baeloha, S. Shoob, O. Scharf, L.R. Heim, N. Buchbinder, O. Shinikamin, I. Shapira, B. Styr, G. Braun, M. Harel, A. Sheinin, N. Geva, Y. Sela, T. Saito, T. Saido, T. Geiger, Y. Nir, Y. Ziv, I. Slutsky, Disrupted neural correlates of anesthesia and sleep reveal early circuit dysfunctions in Alzheimer models. Cell Reports. **38,** 1-25 (2022).
Image Source: https://www.healthywomen.org/created-with-support/important-questions-ask-about-alzheimers-disease

Tiny and Mighty: Featherwing Beetles Demonstrate a New Way to Fly

Eliana Bohn

Miniature featherwing beetles, also called *Paratuposa placentis*, are a feat of nature. Being less than half a millimeter in length, this beetle proves that small can truly be mighty as they are capable of flying at speeds and accelerations of insects three times their size! To better understand the structural mechanisms that allow for this "superpower," scientists created 3D simulations of the insect using precise measurements obtained from a combination of microscopy, videography, and computational methods. The unusual challenges of flight at this scale uncovered a new way to fly: the beetle's bristled wings clap and move like a figure eight to propel forward. Zooming in, its aerodynamics act a bit differently to allow this beetle to metaphorically swim through the air as if it was water. By understanding the mechanics of miniature organisms, learning from nature is a critical source of inspiration in the ever-expanding need and exploration of microscopic industrial products.

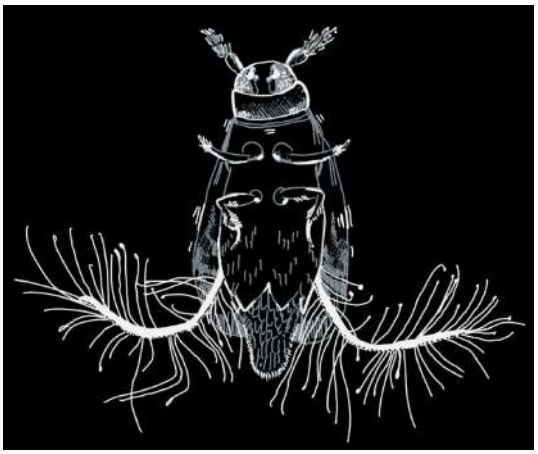


Image: An artistic interpretation of the appearance of *Paratuposa placentis* under a microscope.
Reference: Reference: Sergey E. Farisenkov, et al, Novel flight style and light wings boost flight performance of tiny beetles. Nature. **602**, 96-100 (2022), doi:10.1038/s41586-021-04303-7.
Image Source: Digital illustration by Eliana Bohn.

Biodegradable Banana Paper Can Protect Potatoes from Parasites

Joyce Lui

Potatoes are an important staple for people worldwide. However, potato cyst nematodes (PCNs), a worm-like parasite that attacks potato roots, have been a major pest for farmers globally. Certain synthetic pesticides such as abamectin (ABM) effectively kill developed PCNs, but PCN eggs are protected by a hardened shell that naturally resists these pesticides. To target PCNs over their entire lifespan, researchers from Kenyatta University and multiple other international institutions propose wrapping potatoes with ABM-infused banana paper before planting. Their results indicate that this wrap-and-plant technology not only prevents PCN eggs from hatching, but also allows for slow and sustained release of ABM over weeks, permits healthy root growth, and minimizes environmental contamination. This biodegradable, pesticide-infused banana paper shows potential for ending the PCN plague, as well as similar plagues of plant parasitic nematodes in yams and tomatoes.

Image: Potato cyst nematodes affect potatoes (such as the red potatoes depicted here) globally.
References: J. Ochola et al., Wrap-and-plant technology to manage sustainably potato cyst nematodes in East Africa. Nat Sustain. (2022). doi: 10.1038/s41893-022-00852-5.
Image Source: "Harvesting Potatoes" by FLICKR user Chiot's Run. Licensed under CC BY 2.0. https://www.flickr.com/photos/53255320@N07/6762195603.

AI-Generated Faces Become Increasingly Harder to Detect

Alex Rayo

Artificial intelligence (AI)-generated images, known as deepfakes, are becoming increasingly indistinguishable from real ones. In a series of three experiments, researchers at the University of California, Berkeley tasked participants with classifying 128 images of human faces as real or fake, taken from a set of 400 real and 400 synthetic faces. Other participants were tasked with rating the trustworthiness of 128 faces taken from the same set. The study found that the majority of participants were only slightly better at identifying the AI-generated faces than real ones, and even tended to perceive the synthetic images as more trustworthy than their real counterparts. Consequently, these results suggest that synthetically-generated images are beginning to surpass the uncanny valley. As humans increasingly fail to identify these AI-derived images, concerns over the use of deepface technology for more sinister purposes such as fraud and disinformation campaigns are inevitably raised.

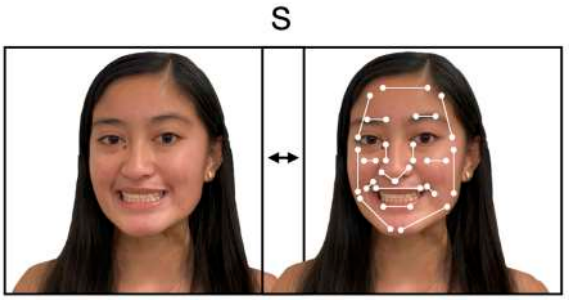


Image: Representation of a synthetic (S) face generated by artificial intelligence.
Reference: S. J. Nightingale, H. Farid, AI-synthesized faces are indistinguishable from real faces and more trustworthy. PNAS. **119,** 8 (2022). https://doi.org/10.1073/pnas.2120481119
Image Source: Digital Illustration by Alex Rayo.

Taming the Untamable: 3D Spatial Control Over AI Image Generation

Kuan Heng (Jordan) Lin

Current AI image generation models can feel like looking into a black box; we often have no idea how they function. 2D models, in particular, have grown exponentially more complex and more difficult to manipulate as they try to capture 3D scenes—but why limit ourselves to 2D? In response to this challenge, Tübingen researchers developed GIRAFFE, a deep neural network model that disentangles objects from the original 2D image and represents them as 3D mathematical encodings called "neural feature fields." GIRAFFE provides complete control over the positions, rotations, and appearances of both the camera view and individual objects with unprecedented accuracy, while having merely a fraction of the network parameters of its predecessors. This research not only strengthens the viability of image generation models in the film, animation, and video game industries but also pushes the boundaries of our artistic control and technical understanding of AI. More than ever, we are closer to taming the untamable.

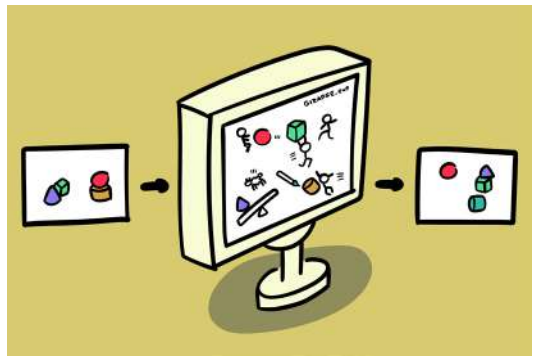


Image: (Extremely) artistic rendition of object translation, rotation, and appearance change with GIRAFFE.
Reference: M. Niemeyer et al. GIRAFFE: Representing scenes as compositional generate neural feature fields. Proc. CVPR IEEE. 11448–11459 (2021).
Image Source: Digital illustration by Jordan Lin.

What Doesn't Kill You Makes You… Smarter? Gut Viruses Linked to Cognition

Caitlyn Brahim

The relationship between the gut microbiome and brain function is primarily studied through gut bacteria. However, recent discovery of a correlation between certain bacteria-infecting viruses in the gut and improved learning ability in animals suggests that bacteriophages also contribute to the gut-brain relationship. To examine this correlation, researchers performed memory tests on humans and mice, as well as examined fruit flies' learning ability to avoid bitter foods. Humans and mice with higher levels of the bacteriophage order *Caudovirales*, particularly the family *Siphoviridae*, and fruit flies fed a whey powder diet rich in *Siphoviridae* demonstrated faster performance than their controls. Researchers concluded that these results support a positive correlation between *Siphoviridae* and the ability to learn and retain information. While a causal role has yet to be established, this study highlights the possibility of using gut-residing bacteriophages in future brain therapies and treatments for cognitive diseases like Alzheimer's.

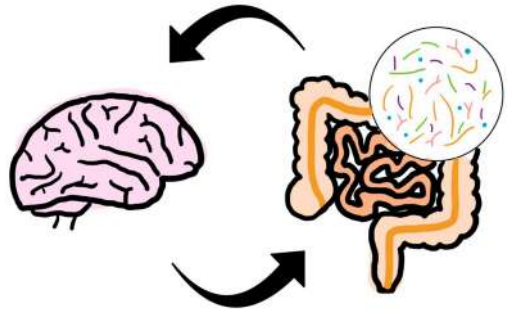


Image: Illustrated schematic of relationship between brain, gut, and gut microbiome.
Reference: J. Mayneris-Perxachs et al., Caudovirales bacteriophages are associated with improved executive function and memory in flies, mice, and humans. Cell Host Microbe. (2022).
Image Source: Digital Illustration by Caitlyn Brahim.

Newly Developed Fabric Coating Capable of Killing SARS-CoV-2 and Other Deadly Pathogens

Chris La

Researchers at the University of British Columbia have developed an inexpensive, non-toxic, antimicrobial coating that decreases the infectivity of the novel coronavirus (SARS-CoV-2) and other deadly pathogens. Companies are already looking to apply the technology to hospital fabrics and activewear. The polymer coating has both passive contact-based and active photodynamic properties, with the latter mode relying on the light-driven generation of reactive oxygen species to damage microbial cells. Modest decreases in viable bacteria count of Escherichia coli and methicillin-resistant Staphylococcus aureus, both major sources of hospital-acquired infections, were observed on infected textiles upon contact with the polymer. A more drastic decrease was observed when the treated textile was exposed to green light. In testing the polymer's efficacy against SARS-CoV-2, the contact-based mode proved ineffective whereas the photodynamic mode resulted in a 90% decrease in infectivity. Unlike other pathogen-fighting technologies, the coating can be used on almost any fabric with a relatively easy and affordable application method.

Image: A white T-shirt and pair of black socks.
Reference: T. Wright et al., Photodynamic and Contact Killing Polymeric Fabric Coating for Bacteria and SARS-CoV-2. ACS Appl. Mater. Inter. **14**, 49–56 (2022). doi: 10.1021/acsami.1c14178
Image Source: Photograph by Pexels user Ryutaro Tsukata. Available under public domain. https://www.pexels.com/photo/set-of-classic-t-shirt-and-socks-for-man-5746086/

Planetary Engulfment Events Said to Happen in a Quarter of Binary Star Systems

Andy He

Binary stars are widely thought to have chemical compositions identical to their companion, as they form from the same material. However, unique binary star systems with two non-identical, inhomogeneous chemically-composed members have been observed, a result that has long puzzled scientists. Recently, Spina et. al, has found compelling evidence indicating that these chemical inhomogeneities may be caused by planetary engulfment events, which involve the assimilation of a planet into the outer layers of a star. These events generally contribute rock-forming elements, such as iron to the compositions of these stars. Spina et. al. arrived at the probability for these planetary engulfment events to happen in binary star systems to be 20-35%. Thus, it is likely that these events cause the chemical structures of binary members to be altered.

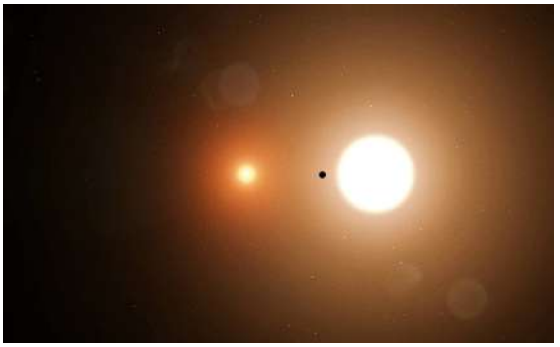


Image: An illustration of a binary star system.
Reference: Spina, L., Sharma, P., Meléndez, J. et al. Chemical evidence for planetary ingestion in a quarter of Sun-like stars. Nat Astron 5, 1163–1169 (2021). doi#10.1038
Image Source: NASA/Goddard Space Flight Center

CRISPR-Cas9 Gene Editing Deployed In Vivo for Disease

Sohan Talluri

Transthyretin amyloidosis is a rare but life-threatening disease that results from the accumulation of misfolded transthyretin (TTR) protein in the heart and nerves. The etiology of this condition lies solely in mutations of the TTR gene, making it a prime target for CRISPR-Cas9 gene therapy. Researchers at Intellia Therapeutics developed a novel Cas9 and lipid nanoparticle therapeutic agent that was infused directly into patients' bloodstreams to knockout the mutated TTR gene, marking the first *in vivo* deployment of CRISPR-Cas9 in a clinical trial. Of the six transthyretin amyloidosis patients treated, the high dose group had an 87% reduction in bloodstream TTR protein with only mild side effects, demonstrating the safety and efficacy of the treatment. Though further clinical trials are needed before generalized application, these results indicate that future *in vivo* CRISPR-Cas9 therapies can safely combat and eliminate monogenic diseases.

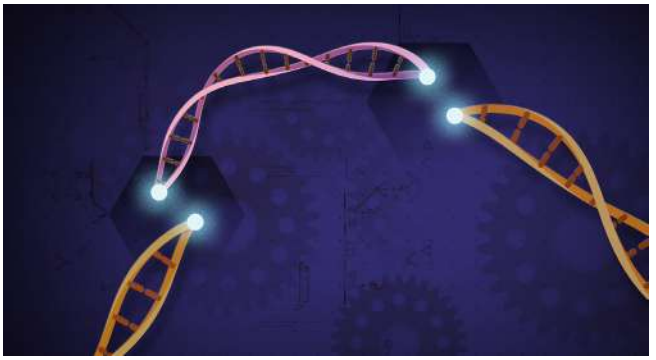


Image: CRISPR-Cas9 cuts double-stranded DNA at specific target sequences.
Reference: J. D. Gillmore *et al.*, CRISPR-Cas9 In Vivo Gene Editing for Transthyretin Amyloidosis. *N. Engl. J. Med.* **385**, 493–502 (2021). doi:10.1056/NEJMoa2107454
Image Source: CRISPR Cas9 by Ernesto del Aguila III, National Human Genome Research Institute, NIH. https://www.flickr.com/photos/nihgov/41124064215. Available under public domain.

How Might a Mushroom Tap Away Your Depression?

Jovian Cheung

Psilocybin, a psychedelic compound derived from 'magic mushrooms', has drawn public attention with its putative ability to lift depressive fogs and conduce mental clarity. In order to examine this potential, researchers at Yale University utilized a rodent model where they tested the compound's effects on mouse behavior after stress and imaged structural changes in cortical neurons during treatment. They observed a visible increase in markers of neuronal connections, which appeared to be partially separate from the drug's known serotonergic mechanism. In addition, electrophysiological recordings of brain slices showed an increase in excitatory cortical activity. These findings provide leads towards how psilocybin exerts its antidepressant actions, as well as implications towards the neurobiological changes underlying depressive disorders. Nevertheless, more work is pending before researchers can develop a safe and effective way of utilizing psilocybin as psychiatric medication - do not try it at home!

Image: Mushrooms that produce the compound psilocybin are colloquially known as "magic mushrooms"
Reference: Ling-Xiao Shao et al., Psilocybin induces rapid and persistent growth of dendritic spines in frontal cortex in vivo. Neuron **109(16)**, 2535-2544.e4 (2021). https://doi.org/10.1016/j.neuron.2021.06.008
Image Source: "Magic mushrooms" by Flickr user kooikkari. Licensed under CC BY-SA 2.0. https://www.flickr.com/photos/126392958@N03/28306445805/

The double-edged implications of cellular senescence on cancer progression

Hieu Nguyen[1,2], Nicole Lynn[2], Emily Martinez[2], Jeffrey Huynh[2], and Jorge Z. Torres[2,3]

[1]Department of Molecular, Cell, and Developmental Biology. [2]Department of Chemistry and Biochemistry. [3]Jonsson Comprehensive Cancer Center.

ABSTRACT

Cellular senescence is an indefinite cell cycle arrest that not only occurs naturally due to telomere shortening, but also activates in response to DNA damage caused by radiation, the accumulation of reactive oxygen species, mitochondrial dysfunction, or other genotoxic anticancer therapies. While this initial defense mechanism limits a damaged cell's potential to become malignant, the extended presence of senescent cells results in the formation of a highly complex secretory microenvironment that can act both to promote and inhibit cancer progression. To clarify the complexities behind senescence and its impact on tumor formation, this review aims to discuss the molecular pathways responsible for establishing senescence, examine the numerous secretory phenotypes that emerge from long-term senescence, and delve into the efficacy and development of current therapeutics used to target senescent behavior.

INTRODUCTION

Cellular senescence, initially defined as an irreversible exit from the cell cycle, was discovered by Leonard Hayflick and Paul Moorhead when they noticed that lung fibroblasts ceased to proliferate after a finite number of cell divisions (*1*). While this definition of senescence was first linked to a progressive decline in cellular regeneration during aging (*2*) due to continual telomere shortening (*3, 4*), an acute activation of senescence can be triggered by oncogene expression and anticancer therapies. Radiation (*5, 6*), the accumulation of reactive oxygen species (ROS) (*7–10*), mitochondrial dysfunction (*11–13*), and genotoxic anticancer therapeutics (*14, 15*) all compromise genomic integrity and lead to the onset of senescence via DNA damage response (DDR) pathway activation (*16*). To acutely direct a cell towards senescence, the DDR initiates the accumulation of tumor suppressor pathways such as the p53/p21 (*17, 18*) and p16-retinoblastoma protein (Rb) (*19*) pathways. These proteins, which operate as cyclin-dependent kinase (CDK) inhibitors, direct the cell towards a G1/S or G2/M cell cycle arrest (*17-19*).

In addition to inhibiting mitotic division, senescence can promote the expression of secreted factors that chronically impact the extracellular environment over time. This collection of secreted factors, also known as the senescence-associated secretory phenotype (SASP), acts as a "double-edged sword" since it can induce both beneficial and harmful cellular changes during tumor development. p53 stabilization during the establishment of cellular senescence triggers SASP formation via activation of the p65 subunit of nuclear factor kappa-B (NF-kB) (*20*). NF-kB then acts as a transcriptional factor, encouraging the expression and secretion of a multitude of proteins including chemokines, interleukins, metalloproteases, and growth factors (*21*). While initial SASP activation can encourage anticancer phenotypes such as the formation of an immune surveillance system that eliminates damaged cells (*22*), the continued maintenance of SASP factor expression by chronic proinflammatory cytokines (*23*) can result in detrimental changes in cell metabolism (*24–26*), angiogenesis (*27, 28*), and proliferation (*29, 30*); these processes assist in cancer progression and metastasis (*31*) (Figure 1).

In order to improve the efficacy of cancer therapeutics that guide cells toward senescence (*32–34*), multiple classes of senotherapeutics are being developed to eradicate the malignant phenotypes that result from chronic senescence. While senomorphic compounds attempt to eliminate these tumorigenic environments by reprogramming or suppressing SASP activity (*35*), other drug classes completely eliminate the existence of senescent cells. Senolytics utilize apoptotic pathways to target senescent cells for destruction (*36*) and DDR inhibitors prevent cells from being prompted towards senescence (*37*). These differences in senotherapies, which are still in the early stages of development, present an ongoing debate on how senotherapeutics should tackle chronic senescence. Is it possible to successfully target long-term oncogenic behaviors of senescent cells, or is it more efficacious to remove them entirely?

Since the complex relationship between senescence and its effects on a tumorigenic environment still remains ambiguous,

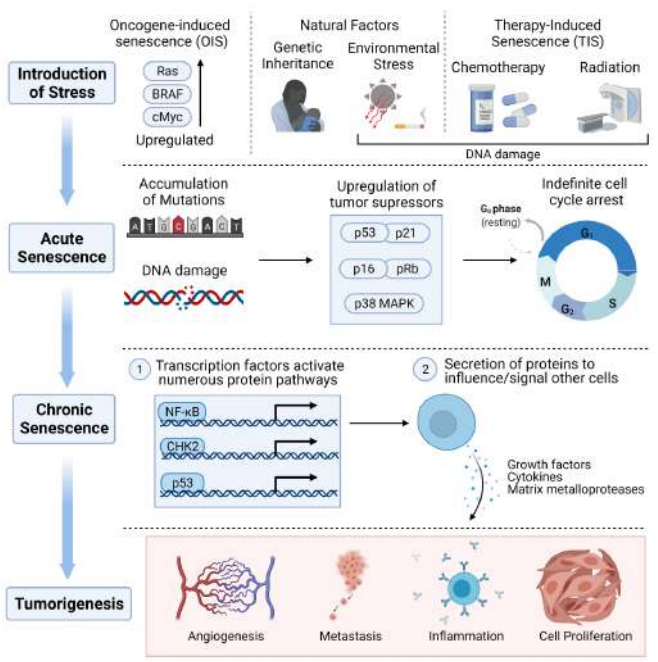


Figure 1. Progression of DNA-damage induced senescence. To limit DNA damage that occurs due to intracellular and extracellular stressors, tumor suppressors guide cells towards senescence, where cells indefinitely remain in G0 phase. Cells arrested in senescence can then impact their extracellular environment through cells activating transcription factors that result in the production and secretion of proteins. The continual secretion of these factors can result in a microenvironment that supports tumorigenesis.

this review aims to clarify the role that senescence plays in both assisting and impeding cancer development. This review also introduces the molecular pathways involved in senescence establishment and SASP development, classifies senotherapeutics and their specific uses, and discusses roadblocks that prevent their functionality. Lastly, this review will identify understudied topics in cancer-related senescence which can bolster the scientific community's understanding of the field through further investigation.

ONCOGENE-INDUCED SENESCENCE

As gene mutations gradually accumulate in the body, the dysregulation of oncogenes, such as Ras and BRAF, leads to the onset of many cancers (*38–40*). To stop these oncogenes from encouraging premalignant behavior, tumor suppressors may guide cells towards oncogene-induced senescence (OIS) as a tumor suppressive mechanism (*18*) (Figure 2). To enter OIS, tumor suppressors, such as p53, p21, p16, and Rb, are upregulated in malignant cells, resulting in the downstream inhibition of cyclin-dependent kinases (CDKs) (*18*). CDKs are then unable to phosphorylate cyclins, which subsequently ceases mitotic progression (*41*). Thus, the accumulation of CDK inhibitors (CDKi) prompts a cell towards G1/S or G2/M cell cycle arrest, allowing cells to transition towards OIS. In oncogenic cells where tumor suppressor pathways are active, p53/p21 regulates cell cycle homeostasis to direct cells towards senescence (*42–44*). While p53 levels are low in healthy, proliferating cells, the activation of oncogenes, such as Myc, Ras, and BRAF, lead to p53 upregulation (*45*, *46*). p53 then transiently activates p21, which binds and inhibits CDK1, CDK2, and CDK4 (*47*), preventing the phosphorylation of cyclins. This guides cells towards a cell cycle arrest: while CDK1 inhibition directs cells to exit at G2/M, inhibition of CDK2 or CDK4 directs cells towards a G1/S exit.

Whereas the p53-p21 pathway activates senescent cell arrest, the p16-Rb tumor suppressor pathway plays a key role in both activating and maintaining OIS (*48*). As p21 expression decreases after the onset of senescence, p16 is activated, preserving the p53-p21 mediated inhibition of CDK1 and CDK2 through a series of inhibitory mechanisms. p16 activation inhibits CDK4 and CDK6 (*49*), which then function to inhibit the retinoblastoma (Rb) protein. As a result, Rb limits cell proliferation by binding to and effectively blocking E2F1 transcription factors, preventing the cell from entering the S phase and redirecting the cell towards OIS at the G1/S transition (*48*, *50*). This is further supported by latent membrane protein 1 (LMP1) suppression of the p16 pathway, promoting cell proliferation and inhibiting OIS (51).

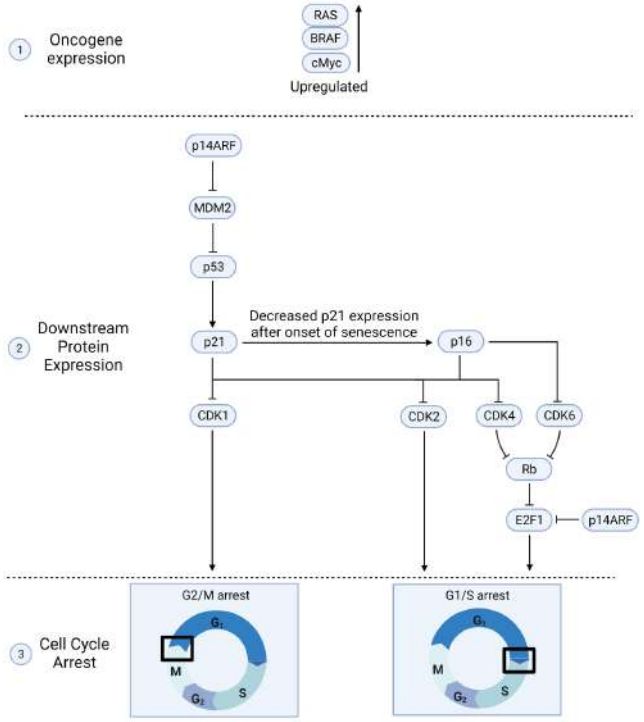


Figure 2. Protein pathways involved in oncogene-induced senescence. To counteract oncogene expression, tumor suppressors will redirect cells to stop proliferation in order to preserve genomic integrity. Using downstream protein expression via the p53-p21 and p16-Rb pathways, cells are redirected either to a G2/M or G1/S cell cycle arrest via inhibition and dephosphorylation of cyclin-dependent kinases.

Abbreviations: CDK, cyclin dependent kinase; G2/M, G2/mitosis; G1/S, G1/synthesis; Rb, retinoblastoma protein; E2F1, E2F Transcription Factor 1.

To control levels of p53/p21, p14ARF directly regulates these tumor suppressors to modulate OIS activation. P14ARF, an isoform of the p16-containing CDKN2A locus, is induced in response to Ras expression (*52*) and maintains p53 levels by preventing degradation by the ubiquitin ligase MDM2. p53 also downregulates p14ARF, creating a negative feedback loop that stabilizes protein levels of p53, p14ARF, and MDM2. The presence of this pathway has been further supported by experiments where the loss of p14ARF promotes MDM2 and causes p53 loss of function (*53*). Finally, p14ARF also binds to the adenosine diphosphate-ribosylation factor (ARF) promoter to downregulate E2F1 transcriptional activity, supplementing the tumor suppressive activity of the p16/Rb pathway.

THERAPY-INDUCED SENESCENCE

Cancer therapeutics such as cytotoxic drugs and radiotherapy focus on inhibiting the rapid proliferation of malignant cells by targeting them for apoptotic death. Many of these therapies purposefully initiate DNA damage (*14*), leading to genotoxic activation of the DDR pathway to prevent further proliferation of oncogenic cells. However, some therapy-driven DDR activation can unintentionally guide a subset of the cell population towards senescence rather than apoptosis (*54*), potentially jeopardizing the efficacy of these treatments due to the development of chronic senescent phenotypes. This process by which cancer therapies halt proliferation through cellular senescence is known as therapy-induced senescence (TIS). Since the classification of senescence-inducing drugs has been discussed previously (*54*), this section will instead examine drugs targeting cell cycle regulators and DNA damaging agents. Through highlighting these therapies, a greater awareness of TIS activation in cancer treatment may be better understood. Chemotherapeutics that target cell cycle regulators are commonly used to cease the progression of cell division in cancerous tissue. Examples of this category include CDK 4/6 inhibitors and microtubule inhibitors (*55*). Firstly, CDK 4/6 inhibitors can block cyclin activity, resulting in a G1/S cell cycle halt. The use of two CDK 4/6 inhibitors, abemaciclib and palbociclib, causes human primary fibroblasts to enter senescence through a progressive loss of Rb phosphorylation and elevated p53 transcriptional levels (*56*). Furthermore, palbociclib treatment induces senescence in esophageal squamous carcinoma cells by inhibiting FOXM1, a transcription factor downstream of CDK 4/6 responsible for cell cycle-dependent gene expression (*57*). Rather than inhibiting protein regulation, microtubule inhibitors impede cell cycle progression by physically destabilizing microtubules and suppressing microtubule dynamics (*56*). Disorazole C1, a common microtubule inhibitor, irreversibly binds to tubulin and prevents polymerization, which results in misaligned chromosomes during mitosis and directs cells towards senescence via a G2/M arrest (*58*). This mitotic exit towards senescence differs from CDK 4/6 inhibitors, which induce a senescent exit during the G1/S transition. Besides targeting the cell cycle, many chemotherapeutics directly damage DNA to induce senescence (*59*). These DNA-damaging agents can be further categorized into forced DNA damage stressors and DNA replication/repair effectors. Forced DNA damage stressors, such as ionizing radiation and alkylating agents, disrupt the chemical structure of DNA (*59*). For example, ionizing radiation causes single-stranded and double-stranded DNA breaks, which signals the p53 dependent pathway in breast tumor cells and promotes TIS (*60*). Exposure to radiation also induces senescence by virtue of increased glycolytic metabolism through the activation of 5'-adenosine monophosphate-activated protein kinase (AMPK). Senescence can also be induced through alkylating agents such as cyclophosphamide (CTX), which inhibits proliferation in lymphoma cells through activation of p53 and p16 (*61*). Lastly, DNA replication/repair effectors disrupt pathways that directly synthesize and maintain DNA. Topoisomerase inhibitors and poly(ADP-ribose) polymerase (PARP) inhibitors are two examples of DNA replication/repair effectors (*62*, *63*). Usage of doxorubicin, a common chemotherapeutic and topoisomerase II inhibitor (*64*), induced TIS through p53-dependent activation of p21 in human colorectal carcinoma cells (*65*). Furthermore, high-grade serous ovarian cancer (HGSOC) cells treated with a common PARP inhibitor (PARPi) induced senescence via the p21 and checkpoint kinase 2 (CHK2) pathways (*66*).

SENESCENCE-ASSOCIATED SECRETORY PHENOTYPE

In addition to indefinitely exiting the active cell cycle, senescent cells retain limited functionality through the SASP, a vast profile of soluble proteins including interleukins, growth factors, and metalloproteases (*67*). Encompassing hundreds of proteins, the SASP proteome is expansive and its variable expression of proteins results in a myriad of phenotypes that redefine cell activity within the tissue microenvironment (*68*). However, this complex regulation of the SASP can paradoxically function to both promote and inhibit tumor formation (Figure 3). Expression of interleukins and chemokines evokes an immune surveillance system that prevents malignancy by removing damaged cells (*22*), yet extended SASP activation has also been shown to result in cell reprogramming that encourages a wide range of tumorigenic behaviors (*24–28*, *31*). SASP establishment has even allowed cells to convert back to proliferation, enabling the reintroduction of cancer into previously treated tissue (*29*, *30*). Thus, the contradictory nature of the SASP raises ambiguity on whether directing cells towards senescence is beneficial towards cancer development. This section will first define the main components of the SASP before delving into the variety of beneficial and harmful SASP-related phenotypes that impact tumor progression.

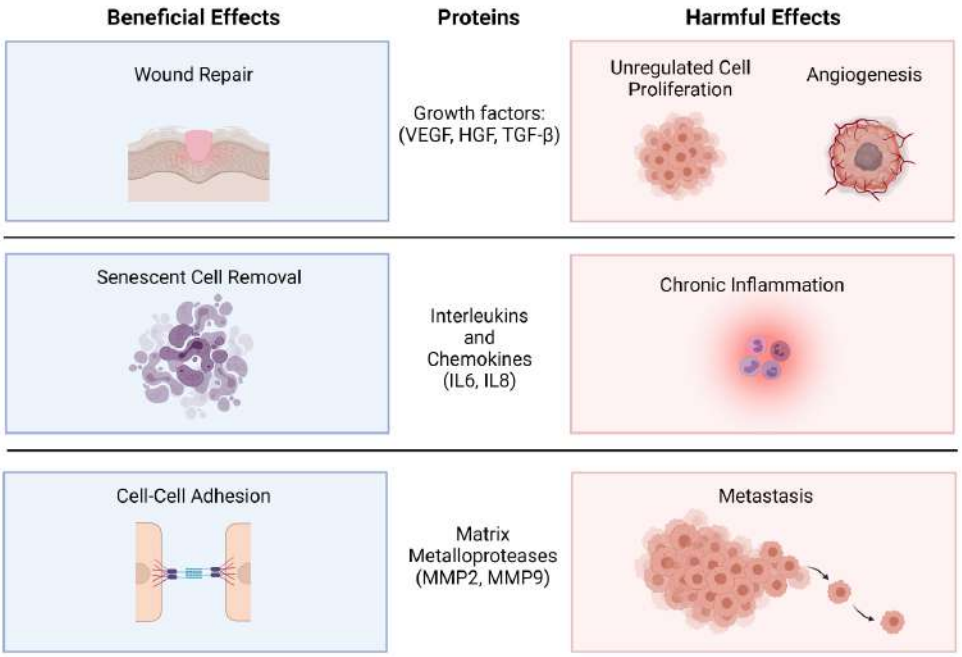


Figure 3. Functions of SASP proteins can both promote and inhibit tumor formation. Although SASP factors initially act to maintain physiological and cellular homeostasis, continual development of the SASP can reprogram these proteins to encourage tumorigenesis. Growth factors involved in wound repair can alter function to encourage cell proliferation and angiogenesis, pro-inflammatory molecules responsible for senescent cell clearance can cause chronic inflammation, and matrix metalloproteinases necessary for cell-cell adhesion can instead promote cell metastasis and cancer invasion.

Abbreviations: VEGF, vascular endothelial growth factor; HGF, hepatocyte growth factor; TGF-β, transforming growth factor beta; IL, interleukin, MMP, matrix metalloproteinase.

SASP ACTIVATION

Formation of the SASP is initiated through protein pathways that respond to various sources of cellular stress. Upon extensive DNA damage, SASP activity is most commonly triggered through ataxia-telangiectasia mutated (ATM) and ATR (ATM and Rad-3 related (ATR) (*69*). When DNA is damaged, these kinases act as primary responders to double-stranded breaks (DSBs), activating downstream transcription factors such as CHK2, NF-kB, and p53 (*70, 71*). These SASP effectors can then act as transcription factors to activate numerous other proteins, making unique proteomes that differ based on the primary stressor, initial genetic profile, cell type, and duration of the SASP (*68*). Despite the diversity of these proteomic profiles, SASP activation commonly leads to the expression of cytokines (IL1α, IL-6, IL-8, CXC chemokines; *72–75*), growth factors (TGF-β, HGF, VEGF; *27, 76, 77*) and ECM-remodeling metalloproteinases (MMP1 and MMP3; *78, 79*).Independent from the canonical DDR pathway's function in directing SASP activity, other protein pathways have been shown to encourage SASP development. For instance, the continual presence of genotoxic and oncogenic stress results in extended phosphorylation of p38 mitogen-activated protein kinases (MAPK), which then promote NF-kB transcriptional activity of SASP factors (*80, 81*). Constant ROS exposure in senescent human endometrium-derived mesenchymal stem cells (hMESCs) for five days resulted in decreased expression of p53/p21 and increased p38 accumulation and phosphorylation (*82*). These results suggest that a DDR-related response is necessary for an acute activation of senescence, while prolonged p38 expression is essential for SASP activity. In addition to the p38 protein pathway, recent research has introduced the potential of other DDR-independent pathways involved in SASP development (*83*). Treatment of cells with a histone deacetylase inhibitor suggested that the accumulation, but not kinase activity, of ATM is necessary for SASP activity in DDR-independent senescence (*83*). Thus, these studies demonstrate that multiple protein pathways are responsible for defining SASP formation and progress.

Beneficial effects of senescence

Immune clearance of pre-malignant cells

One of the primary beneficial functions of the SASP is to identify and remove pre-cancerous cells through an NF-kB-mediated immune surveillance system. While entering senescence, the NF-kB p65 (RelA) subunit accumulates on chromatin, acting as a transcription factor of multiple genes that are essential for senescence-associated phenotypes (74). Notably, NF-kB facilitates the activation of an immune surveillance system by promoting expression of interleukins IL-6 and IL-8, chemokines in the CXCL family, and T-cell activators such as ICAM-1 (84). Activating these proinflammatory molecules spurs the recruitment of an adaptive immune response which seeks to remove cells displaying oncogenic behavior. In mouse livers, post-senescence cytokine release and tissue inflammation

resulted in a CD4+ T-cell-dependent clearance of oncogenic $Nras^{G12V}$ senescent hepatocytes (85). Additionally, inhibiting immune surveillance with perforin, a pore forming protein, resulted in chronic inflammation and a larger abundance of senescent cells (*86*). Thus, senescent activation of an immune surveillance system plays a proactive role in preventing the formation of new malignant cells.

Wound repair

Furthermore, the process of wound healing also drives senescent cell formation and preserves tissue integrity. During wound healing in human fibroblasts, the ECM-associated protein CCN1, also known as cysteine-rich angiogenic protein 61 or CYR61, drives cells into a DDR-directed senescence through the upregulation of p53/p16 and ROS accumulation, resulting in the inhibition of extended fibrosis that can impair tissue integrity and function (*87*). Demaria et al. furthered these studies by demonstrating that the elimination of p16-upregulated senescent mouse embryonic fibroblast (MEF) cells resulted in delayed wound healing and that the secretion of the SASP factor platelet-derived growth factor subunit A (PDGF-A) was responsible for enhancing wound contraction by encouraging myoblast differentiation. Along with growth factors such as PDGF-A, SASP activity also drives the presence of MMP2 and MMP9, which are collagenases that encourage keratinocyte migration that allows for wound closing (88–90). Through the proper regulation of genes involved in ECM function, SASP factors can encourage proactive wound repair mechanisms that positively impact the cell microenvironment. However, reprogrammed expression of proteins involved in both the wound repair and immune surveillance pathways can promote an ideal microenvironment for tumor metastasis and growth, making the SASP a double-edged sword.

Deleterious effects of senescence

Prolonged immune response and chronic inflammation

While expressing inflammatory markers and activating an adaptive immune response can result in the removal of premalignant cells, the extended SASP recruitment of immune cells can result in chronic inflammation of tissues that enhances tumor growth and malignant behavior (*23*). Consistent expression of proinflammatory interleukins and chemokines highly present in SASP profiles have been well-linked to cancer and tumor-promoting alterations such as cell proliferation, neovascularization, and metastasis (91, 92). In addition to encouraging tumorigenic phenotypes, SASP-initiated chronic inflammation has also been shown to alter immune cell response from inhibiting tumor growth. In a study conducted by Guan and company, senescence induced by prolonged exposure to CDK4/6 inhibitors resulted in SASP-mediated recruitment of Gr-1–positive immune cells that reduced antitumor immunity and allowed for cell proliferation (*93*) by regulating cytokines and growth factors that promote tumorigenesis (*94*). Hence, prolonged SASP expression can alter immune cell responses to ensure tumor growth while also fostering a microenvironment favorable for tumorigenesis.

Metastasis

SASP activity involved in wound healing also contributes to metastasis by promoting the epithelial-mesenchymal transition (EMT) and the invasion of cancerous cells into new environments (*95*). The EMT transition state, which transitions endothelial cells into mesenchymal cells, has been associated with tumor initiation, malignant progression, tumor stemness, tumor cell migration, and metastasis (*95*). Activation of the EMT is a complex process that possesses heterogeneity within a cancer cell population, which gives it the ability to achieve a further stage of differentiation and then metastasis after clonal expansion (*96*). Previously, the SASP has been shown to promote the development of cancer metastasis. Human GROα (CXCL-1), a distinct epithelial growth factor released during SASP secretions, is the factor responsible for stimulating growth of pre-malignant and malignant epithelial cells by disrupting epithelial organization through induction of the EMT (*97*). Supporting this, it has been found that C-X-C-motif ligand (CXCL12) receptors are upregulated in senescent cells in vitro and in vivo in papillary thyroid carcinoma (PTC) tissue (*96*). CXCL12 has been identified in CXCL12/CXCR4 signaling which results in a gradient CXCR4 activation, stimulating the directed migration of cancer cells and their invasiveness, promoting metastasis (*98*). Together, these findings demonstrate the role that the SASP has in metastasis.

Angiogenesis

Angiogenesis, the development of blood vessel formation, is essential to tissue repair and maintenance. However, abnormal overexpression of this mechanism can assist tumor growth and expansion through supplying blood to the tumor (*99*). This process encourages further tumor cell proliferation and migration as a result of SASP factors secreted by senescent fibroblasts and cancer cells. Factors including p16 overexpression, replicative exhaustion, and oncogenic RAS expression all contribute to an increased VEGF production in senescent fibroblasts (*27*). VEGF stimulates tumor vascularization by increasing the permeability, migration, and proliferation of vascular endothelial cells (*100*), which could lead to pancreatic (*101*) and breast cancers (*102*). The migration of human umbilical vein endothelial cells (HUVECs) as a result of increased secretion of CXCL12 (also known as SDF-1) also upregulates angiogenesis in senescent cells (*103*). Although angiogenesis is a regulating mechanism that ensures cellular integrity and function, its overexpression in the SASP can cause detrimental damage by aiding tumor growth and enhancing malignant cell migration within the bloodstream.

SENOTHERAPEUTICS

With the increasing awareness of senescence-related tumorigenic phenotypes, existing therapies and new drug targets are currently being investigated to target senescent cells and limit malignant SASP phenotypes. Senotherapies can be classified as senolytics that remove senescent cells

and senomorphics that modify senescent cell behavior. These compounds, which can be used alongside common DNA-damaging drugs (*104*), prevent the potential of cancer relapse spurred by lingering senescent cells that can reprogram their proliferative state once they are no longer inhibited by therapeutic regimens. In this section, we will highlight individual drugs within each of these classes that are either currently being studied in clinical trials or being used in cancer treatments.

The development of senolytics predominantly focuses on targeting the anti-apoptotic BCL-2 protein family, which has been shown to be upregulated in OIS (*105*) and is essential for chemotherapy resistance (*106*). Treatment with the BCL-2, BCL-W, and BCL-XL inhibitor ABT-737 resulted in elimination of senescent cells (*107*), and its derivative navitoclax (ABT-263) is currently being tested in multiple clinical trials (Table 1). However, ABT-737 treatment was also shown to induce proliferation in aggressive senescent variants (*108*), stressing the need for further examination of BCL-2 inhibitors before common clinical use.

Although senolytic usage emphasizes the elimination of senescent cells, senomorphics, such as metformin and ruxolitinib, prevent SASP activity without the removal of senescent cells. Metformin prevents transcription of SASP genes by stopping translocation of NF-kB to the nucleus (*109*), while ruxolitinib is a competitive inhibitor of JAK1/2, a kinase upstream of signal transducers and activators of transcription (STATs) that regulate cell growth and metastasis (*110*). Both metformin and ruxolitinib (Table 2) are currently being tested in a variety of clinical trials, and have been shown experimentally to synergistically work with other cancer drugs (*111*).

Table 1. Senolytics eliminate senescent cells. In order to limit the malignant phenotypes that result from continued senescent cell activity, senolytics target senescent cells for removal. The table above lists existing senolytic compounds as well as their respective mechanisms and potential use to treat different cancers.

Abbreviations: 17-DMAG, 17-Dimethylaminoethylamino-17-demethoxygeldanamycin; HSP, heat shock protein; BCL-2, B-cell lymphoma 2; ROS, reactive oxygen species; PI3K, Phosphoinositide 3-kinase; M, mitosis, YB, Y-box binding protein.

Drug Name	Mechanism	Cancer Types
17-DMAG	HSP90 inhibitor	Lymphoma, Small Intestine Cancer
Dasatinib	Tyrosine kinase inhibitor	Chronic myeloid leukemia; gastrointestinal (colon/rectal) cancers,
Navitoclax	BCL-2 inhibitor	Melanoma, Small Cell Lung Cancer, Myelofibrosis, Advanced/Metastatic Solid Tumors, Myelomonocytic Leukemia
Panobinostat	Histone deacetylase inhibitor	Multiple Myeloma (FDA-approved), Pontine Glioma, Advanced/Metastatic Cancer
Piperlongumine	Accumulation of ROS	gastric cancer, leukemia, multiple myeloma
Quercetin (Flavonoid)	PI3K inhibitor; G2/M arrest or G1 arrest	prostate, cervical, lung, breast, colon cancers
Curcumin	Unidentified; induces senescent cells shown to prevent liver fibrosis and cancer invasion	Intraepithelial neoplasia, prostate cancer
Fisetin (Flavonoid)	Disrupts Wnt signaling; inhibits YB-1 binding protein	cervical, lung, prostate, colorectal, breast cancers; leukemia, melanoma

FUTURE DIRECTIONS AND PERSPECTIVES

Since its identification over half a century ago, the ongoing study of senescence has unmasked the duality of its involvement in both inhibiting and promoting the progression of cancer. Although senescence, by definition, is an irreversible exit from the cell cycle, its long-term secretory reprogramming enhances the proliferative capacity for both the cell and its extracellular microenvironment. This potential for a cell to revert back to proliferation poses a significant problem for senescence-inducing cancer treatments, since patients may be more prone to secondary relapse and metastasis.

Despite the extent of research placed into understanding senescence, there are still many areas of ambiguity that must be further investigated. First, there is no proper method to distinguish senescent cells under all conditions. Although a variety of senescent markers exist (Table 3), no subset of markers can accurately recognize senescent cells across all cell types and stress inducers. Thus, the inconsistent use of biomarkers to define senescence across different studies presents an ongoing challenge on determining whether experimentation accurately distinguishes senescent cells. Additionally, our understanding of senescence induction centers around p53, DDR-dependent models. Since p53 is null in many cancer lines, identifying novel senescent pathways independent of DDR and p53 is essential to accurately understanding the role of senescence within cancer cells.

Table 2. Senomorphics alter SASP function. Rather than eliminate senescent cell populations, senomorphics seek to alter senescent cell activity to inhibit prolonged development of harmful SASP phenotypes. The table above lists existing senomorphic compounds as well as their respective mechanisms and potential use to treat different cancers.

Abbreviations: NF-kB = Nuclear factor kappa B; AMPK = AMP-activated protein kinase; mTOR = mammalian target of rapamycin; JAK1/2 = janus kinase 1/2

Drug Name	Mechanism	Cancer Types	References
Metformin	Prevents translocation of NF-kB; inhibits mitochondrial complex 1; activates AMPK	Breast cancer, atypical endometrial hyperplasia (endometrial cancer), ovarian cancer, fallopian tube cancer, primary peritoneal cancer, Multiple Myeloma,	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12
Rapamycin	Inhibits mTOR	Bladder cancer, Familial Adenomatous Polyposis	13, 14
Ruxolitinib	JAK1/2 inhibitor, GF inhibitor	Leukemia, Breast Cancer, Ovarian cancer	12, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38

Protein	Function	Expression during Senescence	Encoding Gene	Location (Humans)	PMID (Reference Number)
p16 INK4a	Tumor suppressor protein upregulated in senescence that acts as a CDK inhibitor	Upregulated	INK4a	9p21	21880712
p53	Tumor suppressor protein that plays a diverse role in regulating cell cycle proliferation, and is often nonexistent or not functional in cancers	Upregulated	TP53	17p13	32182711
p21	CDK inhibitor that regulates progression from G1 to S	Upregulated	CDKN1A	6p21.2	30190724
SA-β-Gal	pH-specific hydrolase that converts β-galactosides into monosaccharides in senescent cells	Upregulated	lacZ	3p22.3	17634571
Lamin B1	Intermediate filaments that regulate nuclear structure	Downregulated	LMNB1	5p126	22496421
CXCR2	Chemokine receptor that encourages growth arrest	Upregulated	CXCR2	2p218	18555777
IL-6	Proinflammatory cytokine secreted by senescent cells that promotes SASP activity and expansion	Upregulated	IL6	7p22	18555778
IL-8	Proinflammatory cytokine secreted by senescent cells that promotes SASP activity and expansion	Upregulated	CXCL8	4p73	18555778
IGFBP7	A member of a group of proteins responsible for controlling the availability of insulin-like growth factors which can activate signalling processes.	Upregulated	IGFBP7	4p57	18267069
MCP-1	Proinflammatory chemokine responsible for recruiting monocytes as a SASP response	Upregulated	CCL2	17q11	29290100
Oxylipins	Lipophilic signalling molecules that accumulate during senescence, modify SASP behavior, and are secreted during senolysis	Upregulated	PTGER1, PTGER2, PTGER3	19p13.1, 5p13.1, 1p31.1-1p31.2	33811820

Table 3. Potential biomarkers that can identify senescent cells. Due to the myriad of SASP profiles that can develop from senescent cells, a conclusive set of biomarkers to identify all forms of senescence has not been identified yet. The table above lists proteins that either have been identified across a vast array of SASP profiles or play an important role in guiding cells towards senescence.

Abbreviations: CDK, cyclin dependent kinase; SA, senescence-associated; CXCR, chemokine receptor; IL, interleukin, IGFBP, IGF-binding protein, MCP, methyl-accepting chemotaxis proteins.

Finally, it is crucial to evaluate the efficacy of existing and emerging senotherapeutics that are used to treat cancer. While the testing of both senolytics and senomorphics present a conflicting dilemma on whether to respectively destroy senescent cells or modify their behavior, the continuing development of a definitive senescence model will hopefully allow scientists and medical professionals to safely advance these treatments and improve their use in clinical settings.

REFERENCES

1. L. Hayflick, P. S. Moorhead, The serial cultivation of human diploid cell strains. *Exp Cell Res*. **25**, 585–621 (1961). doi:#10.1016/0014-4827(61)90192-6
2. J. Campisi, Cancer, aging and cellular senescence. *In Vivo*. **14**, 183–188 (2000).
3. M. Z. Levy, R. C. Allsopp, A. B. Futcher, C. W. Greider, C. B. Harley, Telomere end-replication problem and cell aging. *J Mol Biol*. **225**, 951–960 (1992). doi:#10.1016/0022-2836(92)90096-3
4. C. B. Harley, A. B. Futcher, C. W. Greider, Telomeres shorten during ageing of human fibroblasts. *Nature*. **345**, 458–460 (1990). doi:#10.1038/345458a0
5. E.-H. Hong *et al*., Ionizing radiation induces cellular senescence of articular chondrocytes via negative regulation of SIRT1 by p38 kinase. *J Biol Chem*. **285**, 1283–1295 (2010). doi:#10.1074/jbc.M109.058628
6. M. Li, L. You, J. Xue, Y. Lu, Ionizing Radiation-Induced Cellular Senescence in Normal, Non-transformed Cells and the Involved DNA Damage Response: A Mini Review. *Front Pharmacol*. **9**, 522 (2018). doi:#10.3389/fphar.2018.00522
7. M. Ogrunc *et al*., Oncogene-induced reactive oxygen species fuel hyperproliferation and DNA damage response activation. *Cell Death Differ*. **21**, 998–1012 (2014). doi:#10.1038/cdd.2014.16
8. J. F. Passos, T. Von Zglinicki, Oxygen free radicals in cell senescence: are they signal transducers? *Free Radic Res*. **40**, 1277–1283 (2006). doi:#10.1080/10715760600917151
9. A. C. Lee *et al*., Ras proteins induce senescence by altering the intracellular levels of reactive oxygen species. *J Biol Chem*. **274**, 7936–7940 (1999). doi:#10.1074/jbc.274.12.7936
10. S. Macip *et al*., Inhibition of p21-mediated ROS accumulation can rescue p21-induced senescence. *EMBO J*. **21**, 2180–2188 (2002). doi:#10.1093/emboj/21.9.2180
11. P. V. S. Vasileiou *et al*., Mitochondrial Homeostasis and Cellular Senescence. *Cells*. **8**, E686 (2019). doi:#10.3390/cells8070686
12. G. Ye *et al*., Oxidative stress-mediated mitochondrial dysfunction facilitates mesenchymal stem cell senescence in ankylosing spondylitis. *Cell Death Dis*. **11**, 1–13 (2020). doi:#10.1038/s41419-020-02993-x
13. O. Moiseeva, V. Bourdeau, A. Roux, X. Deschênes-Simard, G. Ferbeyre, Mitochondrial dysfunction contributes to oncogene-induced senescence. *Mol Cell Biol*. **29**, 4495–4507 (2009). doi:#10.1128/MCB.01868-08
14. R. H. te Poele, A. L. Okorokov, L. Jardine, J. Cummings, S. P. Joel, DNA damage is able to induce senescence in tumor cells in vitro and in vivo. *Cancer Res*. **62**, 1876–1883 (2002).
15. B. Wang, J. Kohli, M. Demaria, Senescent Cells in Cancer Therapy: Friends or Foes? *Trends Cancer*. **6**, 838–857 (2020). doi:#10.1016/j.trecan.2020.05.004
16. T. von Zglinicki, G. Saretzki, J. Ladhoff, F. d'Adda di Fagagna, S. P. Jackson, Human cell senescence as a DNA damage response. *Mech Ageing Dev*. **126**, 111–117 (2005). doi:#10.1016/j.mad.2004.09.034
17. W. Xue *et al*., Senescence and tumour clearance is triggered by p53 restoration in murine liver carcinomas. *Nature*. **445**, 656–660 (2007). doi:#10.1038/nature05529
18. M. Serrano, A. W. Lin, M. E. McCurrach, D. Beach, S. W. Lowe, Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. *Cell*. **88**, 593–602 (1997). doi:#10.1016/s0092-8674(00)81902-9
19. J. J. L. Jacobs, T. de Lange, Significant role for p16INK4a in p53-independent telomere-directed senescence. *Curr Biol*. **14**, 2302–2308 (2004). doi:#10.1016/j.cub.2004.12.025
20. G. Schneider *et al*., Cross talk between stimulated NF-kappaB and the tumor suppressor p53. *Oncogene*. **29**, 2795–2806 (2010). doi:#10.1038/onc.2010.46
21. A. Salminen, A. Kauppinen, K. Kaarniranta, Emerging role of NF-κB signaling in the induction of senescence-associated secretory phenotype (SASP). *Cell Signal*. **24**, 835–845 (2012). doi:#10.1016/j.cellsig.2011.12.006
22. L. Hoenicke, L. Zender, Immune surveillance of senescent cells--biological significance in cancer- and non-cancer pathologies. *Carcinogenesis*. **33**, 1123–1126 (2012). doi:#10.1093/carcin/bgs124
23. J.-P. Coppé, P.-Y. Desprez, A. Krtolica, J. Campisi, The Senescence-Associated Secretory Phenotype: The Dark Side of Tumor Suppression. *Annu Rev Pathol*. **5**, 99–118 (2010). doi:#10.1146/annurev-pathol-121808-102144
24. M.-L. Kuo *et al*., RRM2B suppresses activation of the oxidative stress pathway and is up-regulated by p53 during senescence. *Sci Rep*. **2**, 822 (2012). doi:#10.1038/srep00822
25. K. M. Aird *et al*., Suppression of nucleotide metabolism underlies the establishment and maintenance of oncogene-induced senescence. *Cell Rep*. **3**, 1252–1265 (2013). doi:#10.1016/j.celrep.2013.03.004

26. E. L. James *et al*., Senescent human fibroblasts show increased glycolysis and redox homeostasis with extracellular metabolomes that overlap with those of irreparable DNA damage, aging, and disease. *J Proteome Res.* **14**, 1854–1871 (2015). doi:#10.1021/pr501221g
27. J.-P. Coppé, K. Kauser, J. Campisi, C. M. Beauséjour, Secretion of Vascular Endothelial Growth Factor by Primary Human Fibroblasts at Senescence *. *Journal of Biological Chemistry*. **281**, 29568–29574 (2006). doi:#10.1074/jbc.M603307200
28. M. Oubaha *et al*., Senescence-associated secretory phenotype contributes to pathological angiogenesis in retinopathy. *Sci Transl Med*. **8**, 362ra144 (2016). doi:#10.1126/scitranslmed.aaf9440
29. Z. V. Chitikova *et al*., Sustained activation of DNA damage response in irradiated apoptosis-resistant cells induces reversible senescence associated with mTOR downregulation and expression of stem cell markers. *Cell Cycle.* **13**, 1424–1439 (2014). doi:#10.4161/cc.28402
30. M. Nakamura, S. Ohsawa, T. Igaki, Mitochondrial defects trigger proliferation of neighbouring cells via a senescence-associated secretory phenotype in Drosophila. *Nat Commun.* **5**, 5264 (2014). doi:#10.1038/ncomms6264
31. I. Guccini *et al*., Senescence Reprogramming by TIMP1 Deficiency Promotes Prostate Cancer Metastasis. *Cancer cell.* **39**, 68-82.e9 (2021). doi:#10.1016/j.ccell.2020.10.012
32. E. Piegari *et al*., Doxorubicin induces senescence and impairs function of human cardiac progenitor cells. *Basic Res Cardiol.* **108**, 334 (2013). doi:#10.1007/s00395-013-0334-4
33. M. Bang, D. G. Kim, E. L. Gonzales, K. J. Kwon, C. Y. Shin, Etoposide Induces Mitochondrial Dysfunction and Cellular Senescence in Primary Cultured Rat Astrocytes. *Biomol Ther (Seoul).* **27**, 530–539 (2019). doi:#10.4062/biomolther.2019.151
34. X. Wang *et al*., Evidence of cisplatin-induced senescent-like growth arrest in nasopharyngeal carcinoma cells. *Cancer Res*. **58**, 5019–5022 (1998).
35. J. S. Lim *et al*., Identification of a novel senomorphic agent, avenanthramide C, via the suppression of the senescence-associated secretory phenotype. *Mech Ageing Dev.* **192**, 111355 (2020). doi:#10.1016/j.mad.2020.111355
36. Y. Zhu *et al*., Identification of a novel senolytic agent, navitoclax, targeting the Bcl-2 family of anti-apoptotic factors. *Aging Cell.* **15**, 428–435 (2016). doi:#10.1111/acel.12445
37. H. T. Kang *et al*., Chemical screening identifies ATM as a target for alleviating senescence. *Nat Chem Biol*. **13**, 616–623 (2017). doi:#10.1038/nchembio.2342
38. A. P. Czernilofsky *et al*., Nucleotide sequence of an avian sarcoma virus oncogene (src) and proposed amino acid sequence for gene product. *Nature.* **287**, 198–203 (1980). doi:#10.1038/287198a0
39. G. S. Martin, Rous sarcoma virus: a function required for the maintenance of the transformed state. *Nature.* **227**, 1021–1023 (1970). doi:#10.1038/2271021a0
40. P. H. Duesberg, P. K. Vogt, RNA species obtained from clonal lines of avian sarcoma and from avian leukosis virus. *Virology*. **54**, 207–219 (1973). doi:#10.1016/0042-6822(73)90130-x
41. F. Fang, J. W. Newport, Evidence that the G1-S and G2-M transitions are controlled by different cdc2 proteins in higher eukaryotes. *Cell.* **66**, 731–742 (1991). doi:#10.1016/0092-8674(91)90117-h
42. S. Modi *et al*., Protein expression of the RB-related gene family and SV40 large T antigen in mesothelioma and lung cancer. *Oncogene.* **19**, 4632–4639 (2000). doi:#10.1038/sj.onc.1203815
43. J. W. Harbour *et al*., Abnormalities in structure and expression of the human retinoblastoma gene in SCLC. *Science*. **241**, 353–357 (1988). doi:#10.1126/science.2838909
44. M. Fischer, M. Quaas, L. Steiner, K. Engeland, The p53-p21-DREAM-CDE/CHR pathway regulates G2/M cell cycle genes. *Nucleic Acids Res*. **44**, 164–174 (2016). doi:#10.1093/nar/gkv927
45. D. W. Felsher, A. Zetterberg, J. Zhu, T. Tlsty, J. M. Bishop, Overexpression of MYC causes p53-dependent G2 arrest of normal fibroblasts. *PNAS*. **97**, 10544–10548 (2000). doi:#10.1073/pnas.190327097
46. D. J. Ahnen *et al*., Ki-ras mutation and p53 overexpression predict the clinical behavior of colorectal cancer: a Southwest Oncology Group study. *Cancer Res*. **58**, 1149–1158 (1998).
47. J. W. Harper, G. R. Adami, N. Wei, K. Keyomarsi, S. J. Elledge, The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. *Cell.* **75**, 805–816 (1993). doi:#10.1016/0092-8674(93)90499-g
48. F. Bunz *et al*., Requirement for p53 and p21 to sustain G2 arrest after DNA damage. *Science*. **282**, 1497–1501 (1998). doi:#10.1126/science.282.5393.1497
49. G. H. Stein, L. F. Drullinger, A. Soulard, V. Dulić, Differential roles for cyclin-dependent kinase inhibitors p21 and p16 in the mechanisms of senescence and differentiation in human fibroblasts. *Mol Cell Biol*. **19**, 2109–2117 (1999). doi:#10.1128/MCB.19.3.2109
50. S. Bagchi, R. Weinmann, P. Raychaudhuri, The retinoblastoma protein copurifies with E2F-I, an E1A-regulated inhibitor of the transcription factor E2F. *Cell.* **65**, 1063–1072 (1991). doi:#10.1016/0092-8674(91)90558-g
51. N. Ohtani *et al*., Epstein-Barr virus LMP1 blocks p16INK4a–RB pathway by promoting nuclear export of E2F4/5. *J Cell Biol*. **162**, 173–183 (2003). doi:#10.1083/jcb.200302085
52. G. Ferbeyre *et al*., PML is induced by oncogenic ras and promotes premature senescence. *Genes Dev*. **14**, 2015–2027 (2000).
53. B. Eymin, S. Gazzeri, C. Brambilla, E. Brambilla, Mdm2 overexpression and p14(ARF) inactivation are two mutually exclusive events in primary human lung tumors. *Oncogene.* **21**, 2750–2761 (2002). doi:#10.1038/sj.onc.1205359
54. J. A. Ewald, J. A. Desotelle, G. Wilding, D. F. Jarrard, Therapy-induced senescence in cancer. *J Natl Cancer Inst*. **102**, 1536–1546 (2010). doi:#10.1093/jnci/djq364
55. M. Demaria *et al*., Cellular Senescence Promotes Adverse Effects of Chemotherapy and Cancer Relapse. *Cancer Discov*. **7**, 165–176 (2017). doi:#10.1158/2159-8290.CD-16-0241
56. S. Goel, M. J. DeCristo, S. S. McAllister, J. J. Zhao, CDK4/6 Inhibition in Cancer: Beyond Cell Cycle Arrest. *Trends in Cell Biology*. **28**, 911–925 (2018). doi:#10.1016/j.tcb.2018.07.002
57. L. Chen, J. Pan, Dual cyclin-dependent kinase 4/6 inhibition by PD-0332991 induces apoptosis and senescence in oesophageal squamous cell carcinoma cells. *Br J Pharmacol*. **174**, 2427–2443 (2017). doi:#10.1111/bph.13836
58. M. B. Tierno *et al*., Microtubule binding and disruption and induction of premature senescence by disorazole C(1). *J Pharmacol Exp Ther*. **328**, 715–722 (2009). doi:#10.1124/jpet.108.147330
59. K. Cheung-Ong, G. Giaever, C. Nislow, DNA-damaging agents in cancer chemotherapy: serendipity and chemical biology. *Chem Biol.* **20**, 648–659 (2013). doi:#10.1016/j.chembiol.2013.04.007
60. K. R. Jones *et al*., p53-Dependent accelerated senescence induced by ionizing radiation in breast tumour cells. *Int J Radiat Biol*. **81**, 445–458 (2005). doi:#10.1080/09553000500168549
61. C. A. Schmitt *et al*., A senescence program controlled by p53 and p16INK4a contributes to the outcome of cancer therapy. *Cell.* **109**, 335–346 (2002). doi:#10.1016/s0092-8674(02)00734-1
62. K. Hevener, T. A. Verstak, K. E. Lutat, D. L. Riggsbee, J. W. Mooney, Recent developments in topoisomerase-targeted cancer chemotherapy. *Acta Pharm Sin B.* **8**, 844–861 (2018). doi:#10.1016/j.apsb.2018.07.008
63. A. Chen, PARP inhibitors: its role in treatment of cancer. *Chin J Cancer*. **30**, 463–471 (2011). doi:#10.5732/cjc.011.10111
64. S. Rivankar, An overview of doxorubicin formulations in cancer therapy. *J Cancer Res Ther*. **10**, 853–858 (2014). doi:#10.4103/0973-1482.139267
65. M. A. Sliwinska *et al*., Induction of senescence with doxorubicin leads to increased genomic instability of HCT116 cells. *Mech Ageing Dev.* **130**, 24–32 (2009). doi:#10.1016/j.mad.2008.04.011
66. K. Bajbouj, J. Shafarin, J. Taneera, M. Hamad, Estrogen Signaling Induces Mitochondrial Dysfunction-Associated Autophagy and Senescence in Breast Cancer Cells. *Biology (Basel).* **9**, 68 (2020). doi:#10.3390/biology9040068
67. S. Watanabe, S. Kawamoto, N. Ohtani, E. Hara, Impact of senescence-associated secretory phenotype and its potential as a therapeutic target for senescence-associated diseases. *Cancer Sci.* **108**, 563–569 (2017). doi:#10.1111/cas.13184
68. N. Basisty *et al*., A proteomic atlas of senescence-associated secretomes for aging biomarker development. *PLoS Biol.* **18**, e3000599 (2020). doi:#10.1371/journal.pbio.3000599
69. A. Maréchal, L. Zou, DNA damage sensing by the ATM and ATR kinases. *Cold Spring Harb Perspect Biol.* **5**, a012716 (2013). doi:#10.1101/cshperspect.a012716

70. J. Zhao *et al*., ATM is a key driver of NF-κB-dependent DNA-damage-induced senescence, stem cell dysfunction and aging. *Aging (Albany NY).* **12**, 4688–4710 (2020). doi:#10.18632/aging.102863
71. J. Y. Lee *et al*., GATA4-dependent regulation of the secretory phenotype via MCP-1 underlies lamin A-mediated human mesenchymal stem cell aging. *Exp Mol Med*. **50**, 1–12 (2018). doi:#10.1038/s12276-018-0092-3
72. C. Xia, J. K. Cheshire, H. Patel, P. Woo, Cross-talk between transcription factors NF-kappa B and C/EBP in the transcriptional regulation of genes. *Int J Biochem Cell Biol*. **29**, 1525–1539 (1997). doi:#10.1016/s1357-2725(97)00083-6
73. R.-M. Laberge *et al*., MTOR regulates the pro-tumorigenic senescence-associated secretory phenotype by promoting IL1A translation. *Nat Cell Biol.* **17**, 1049–1061 (2015). doi:#10.1038/ncb3195
74. Q. Zeng *et al*., Identification of Therapeutic Targets and Prognostic Biomarkers Among CXC Chemokines in the Renal Cell Carcinoma Microenvironment. *Front Oncol*. **9**, 1555 (2019). doi:#10.3389/fonc.2019.01555
75. P. Ortiz-Montero, A. Londoño-Vallejo, J.-P. Vernot, Senescence-associated IL-6 and IL-8 cytokines induce a self- and cross-reinforced senescence/inflammatory milieu strengthening tumorigenic capabilities in the MCF-7 breast cancer cell line. *Cell Commun Signal*. **15**, 17 (2017). doi:#10.1186/s12964-017-0172-3
76. Y. Li *et al*., DNA Damage Activates TGF-β Signaling via ATM-c-Cbl-Mediated Stabilization of the Type II Receptor TβRII. *Cell Rep*. **28**, 735-745.e4 (2019). doi:#10.1016/j.celrep.2019.06.045
77. J. Mikuła-Pietrasik *et al*., Oxidative stress contributes to hepatocyte growth factor-dependent pro-senescence activity of ovarian cancer cells. *Free Radic Biol Med*. **110**, 270–279 (2017). doi:#10.1016/j.freeradbiomed.2017.06.015
78. S. Parrinello, J.-P. Coppe, A. Krtolica, J. Campisi, Stromal-epithelial interactions in aging and cancer: senescent fibroblasts alter epithelial cell differentiation. *J Cell Sci.* **118**, 485–496 (2005). doi:#10.1242/jcs.01635
79. J. Dasgupta *et al*., Reactive oxygen species control senescence-associated matrix metalloproteinase-1 through c-Jun-N-terminal kinase. *J Cell Physiol*. **225**, 52–62 (2010). doi:#10.1002/jcp.22193
80. A. Freund, C. K. Patil, J. Campisi, p38MAPK is a novel DNA damage response-independent regulator of the senescence-associated secretory phenotype. *EMBO J.* **30**, 1536–1548 (2011). doi:#10.1038/emboj.2011.69
81. A. Barascu *et al*., Oxidative stress induces an ATM-independent senescence pathway through p38 MAPK-mediated lamin B1 accumulation. *EMBO J*. **31**, 1080–1094 (2012). doi:#10.1038/emboj.2011.492
82. A. Borodkina, A. Shatrova, P. Abushik, N. Nikolsky, E. Burova, Interaction between ROS dependent DNA damage, mitochondria and p38 MAPK underlies senescence of human adult stem cells. *Aging (Albany NY)*. **6**, 481–495 (2014). doi:#10.18632/aging.100673
83. N. Malaquin *et al*.,Non-canonical ATM/MRN activities temporally define the senescence secretory program. *EMBO Rep*. **21**, e50718 (2020). doi:#10.15252/embr.202050718
84. Y. Chien *et al*., Control of the senescence-associated secretory phenotype by NF-κB promotes senescence and enhances chemosensitivity. *Genes Dev*. **25**, 2125–2136 (2011). doi:#10.1101/gad.17276711
85. T.-W. Kang *et al*., Senescence surveillance of pre-malignant hepatocytes limits liver cancer development. *Nature.* **479**, 547–551 (2011). doi:#10.1038/nature10599
86. Y. Ovadya *et al*., Impaired immune surveillance accelerates accumulation of senescent cells and aging. *Nat Commun*. **9**, 5435 (2018). doi:#10.1038/s41467-018-07825-3
87. M. Demaria *et al*., An essential role for senescent cells in optimal wound healing through secretion of PDGF-AA. *Dev Cell*. **31**, 722–733 (2014). doi:#10.1016/j.devcel.2014.11.012
88. D. Telgenhoff, B. Shroot, Cellular senescence mechanisms in chronic wound healing. *Cell Death Differ*. **12**, 695–698 (2005). doi:#10.1038/sj.cdd.4401632
89. Y. Hassona, N. Cirillo, K. Heesom, E. K. Parkinson, S. S. Prime, Senescent cancer-associated fibroblasts secrete active MMP-2 that promotes keratinocyte dis-cohesion and invasion. *Br J Cancer*. **111**, 1230–1237 (2014). doi:#10.1038/bjc.2014.438
90. Y. Ma *et al*., Deriving a cardiac ageing signature to reveal MMP-9-dependent inflammatory signalling in senescence. *Cardiovasc Res*. **106**, 421–431 (2015). doi:#10.1093/cvr/cvv128
91. M. T. Chow, A. D. Luster, Chemokines in Cancer. *Cancer Immunol Res*. **2**, 1125–1131 (2014). doi:#10.1158/2326-6066.CIR-14-0160
92. L. M. Coussens, Z. Werb, Inflammation and cancer. *Nature.* **420**, 860–867 (2002). doi:#10.1038/nature01322
93. Z. Li *et al*., Gr-1+CD11b+ cells are responsible for tumor promoting effect of TGF-β in breast cancer progression. *Int J Cancer*. **131**, 2584–2595 (2012). doi:#10.1002/ijc.27572
94. H. H. Yan *et al*., Gr-1+CD11b+ myeloid cells tip the balance of immune protection to tumor promotion in the premetastatic lung. *Cancer Res*. **70**, 6139–6149 (2010). doi:#10.1158/0008-5472.CAN-10-0706
95. R.-M. Laberge, P. Awad, J. Campisi, P.-Y. Desprez, Epithelial-mesenchymal transition induced by senescent fibroblasts. *Cancer Microenviron*. **5**, 39–44 (2012). doi:#10.1007/s12307-011-0069-4
96. I. Pastushenko, C. Blanpain, EMT Transition States during Tumor Progression and Metastasis. *Trends Cell Biol*. **29**, 212–226 (2019). doi:#10.1016/j.tcb.2018.12.001
97. Y. H. Kim *et al*., Senescent tumor cells lead the collective invasion in thyroid cancer. *Nat Commun*. **8**, 15208 (2017). doi:#10.1038/ncomms15208
98. S. Heerboth *et al*., EMT and tumor metastasis. *Clin Transl Med*. **4**, 6 (2015). doi:#10.1186/s40169-015-0048-3
99. H. Saman, S. S. Raza, S. Uddin, K. Rasul, Inducing Angiogenesis, a Key Step in Cancer Vascularization, and Treatment Approaches. *Cancers (Basel)*. **12**, 1172 (2020). doi:#10.3390/cancers12051172
100. Y. Watanabe, S. W. Lee, M. Detmar, I. Ajioka, H. F. Dvorak, Vascular permeability factor/vascular endothelial growth factor (VPF/VEGF) delays and induces escape from senescence in human dermal microvascular endothelial cells. *Oncogene.* **14**, 2025–2032 (1997). doi:#10.1038/sj.onc.1201033
101. K. J. Higgins, M. Abdelrahim, S. Liu, K. Yoon, S. Safe, Regulation of vascular endothelial growth factor receptor-2 expression in pancreatic cancer cells by Sp proteins. *Biochem Biophys Res Commun*. **345**, 292–301 (2006). doi:#10.1016/j.bbrc.2006.04.111
102. L. Rydén *et al*., Tumor-specific VEGF-A and VEGFR2 in postmenopausal breast cancer patients with long-term follow-up. Implication of a link between VEGF pathway and tamoxifen response. *Breast Cancer Res Treat*. **89**, 135–143 (2005). doi:#10.1007/s10549-004-1655-7
103. M. Zhang *et al*., CXCL12 enhances angiogenesis through CXCR7 activation in human umbilical vein endothelial cells. *Sci Rep*. **7**, 8289 (2017). doi:#10.1038/s41598-017-08840-y
104. C. J. Sieben, I. Sturmlechner, B. van de Sluis, J. M. van Deursen, Two-Step Senescence-Focused Cancer Therapies. *Trends in Cell Biology*. **28**, 723–737 (2018). doi:#10.1016/j.tcb.2018.04.006
105. B. Tombor, K. Rundell, Z. N. Oltvai, Bcl-2 promotes premature senescence induced by oncogenic Ras. *Biochem Biophys Res Commun*. **303**, 800–807 (2003). doi:#10.1016/s0006-291x(03)00402-9
106. U. A. Sartorius, P. H. Krammer, Upregulation of Bcl-2 is involved in the mediation of chemotherapy resistance in human small cell lung cancer cell lines. *Int J Cancer*. **97**, 584–592 (2002). doi:#10.1002/ijc.10096
107. R. Yosef *et al*., Directed elimination of senescent cells by inhibition of BCL-W and BCL-XL. *Nat Commun*. **7**, 11190 (2016). doi:#10.1038/ncomms11190
108. L. Yang, J. Fang, J. Chen, Tumor cell senescence response produces aggressive variants. *Cell Death Discov*. **3**, 17049 (2017). doi:#10.1038/cddiscovery.2017.49
109. O. Moiseeva *et al*., Metformin inhibits the senescence-associated secretory phenotype by interfering with IKK/NF-κB activation. *Aging Cell*. **12**, 489–498 (2013). doi:#10.1111/acel.12075
110. B. L. Wehde *et al*., Janus Kinase 1 Plays a Critical Role in Mammary Cancer Progression. *Cell Rep*. **25**, 2192-2207.e5 (2018). doi:#10.1016/j.celrep.2018.10.063
111. E. S. Han *et al*., Ruxolitinib synergistically enhances the anti-tumor activity of paclitaxel in human ovarian cancer. *Oncotarget*. **9**, 24304–24319 (2018). doi:#10.18632/oncotarget.24368
112. Clinical Trials Search Results - National Cancer Institute (2016), (available at https://www.cancer.gov/about-cancer/treatment/clinical-trials/search).

Reconstruction of glacial and deglacial hydroclimates at Lake Tauca

Andrea de Guzman-Metivier[1], Alexandrea Arnold[1,2,3], Lauren Santi[3,4], Aradhna Tripati, PhD[1,2,3,4,5]

[1]Department of Atmospheric and Oceanic Sciences. [2]Institute of the Environment and Sustainability. [3]Center for Diverse Leadership in Science. [4]Department of Earth, Planetary, and Space Sciences. [5]School of Earth Sciences, University of Bristol, United Kingdom.

ABSTRACT

During the Last Glacial Maximum (LGM; 23-19 thousand years before present; ka BP), the Bolivian Altiplano region in South America contained many lakes, some of which no longer exist. Lake Tauca occupied the largest closed basin in South America. It is of particular interest due to its sheer size, comparable to some of the largest present-day lakes, and subsequent desiccation to what is now known as Lake Titicaca. Here the relative contributions of three hydroclimate variables - precipitation, temperature, and evaporation - on sustaining Lake Tauca during the LGM and driving the disappearance of the lake during the subsequent deglaciation are tested. Clumped isotope data (heavy isotope pairing of ^{13}C and ^{18}O) for shoreline sediments is reported and the relative contribution of each environmental factor to the lake's hydroclimate is modeled. Preliminary results provide evidence for slightly cooler temperatures during these past intervals relative to modern, with increased precipitation rates and decreased evaporation rates. These results shed light on changes in the water cycle in the water-sensitive Altiplano and provide insight on regional hydrologic conditions under different climate forcings during the Tauca highstand (18 to 14.5 ka BP) and subsequent Coipasa highstand (12.5 to 11.9 ka BP). Results from this study are used to evaluate the TraCE-21ka model hindcasts of climate in South America, showing reasonable agreement with reconstructed temperature but poor agreement in precipitation, indicating a key area for model improvement.

INTRODUCTION

The Last Glacial Maximum (LGM) and the deglaciation that followed (19-11 ka BP) was the last major global climatological transition. During this time, temperatures in the Western Hemisphere increased 4 - 8°C on average and precipitation varied strongly throughout (*1*). In western South America, these intervals were marked by increased moisture and extensive lake systems such as Lake Tauca in the Bolivian Altiplano (Figure 1), many of which reduced in size or disappeared (*2*). Lake Tauca was a closed basin lake (*3*), meaning that the lake had surface water inflows but no outflows, increasing the sensitivity of lake level in the basin to climate fluctuations. Therefore, the primary input into the lake was precipitation (and related surface runoff from the lake's watershed), and the primary output from the lake was evaporation. Lake Tauca's reduction in size to present-day Lake Titicaca, with a change in surface area from ~51,000 to ~9,000 km^2 (*2*), implies it was a significant influence on the regional ecology. This system provides a unique natural laboratory to understand climatic drivers of the water cycle on a regional scale.

During the LGM, the Bolivian Altiplano was wetter primarily due to lower evaporation rates compared to modern. The shift to a drier climate today is attributed to a climate forcing caused by changes in sea surface temperatures (SST), which are associated with El Niño Southern Oscillation (ENSO), a result of Bolivia's geographic location and close proximity to the Pacific Ocean (*4*). These changes in climate conditions are exemplified in paleoclimate proxies such as pollen, glaciers, and shoreline carbonates, which suggest that deglaciation impacted regional climate. Evaluating hydroclimate proxies is useful because these proxies serve as snapshots of the climatic conditions of the region and time in which they are formed, giving us the ability to track climatic changes through time. In addition to a long-term shift in climate, there are several known major short-lived climate excursions during the deglaciation, that may have been associated with changes in precipitation and evaporation, including Heinrich Stadial 1 (HS1; 18.5-14.5 ka BP) which is associated with Northern Hemisphere cooling, the Bølling–Allerød period (BA; 14.7-12.9 ka BP) associated with Northern Hemisphere warming, and the Younger Dryas cooling (YD; 12.9-11.7 ka BP) (*5*).

Climate reconstructions derived from pollen (*6*), glaciers (*7*), and shoreline carbonates (*8*) typically support there having been wetter-than-modern hydroclimates in the region during the LGM, HS1, and YD that are known today. Fluctuations in air and sea surface temperature in the North Atlantic and Greenland during the Late Pleistocene (HS1, BA, YD) are correlated with fluctuations in lake levels. Lake Tauca can be split up into two cycles: the Tauca highstand and the Coipasa highstand.

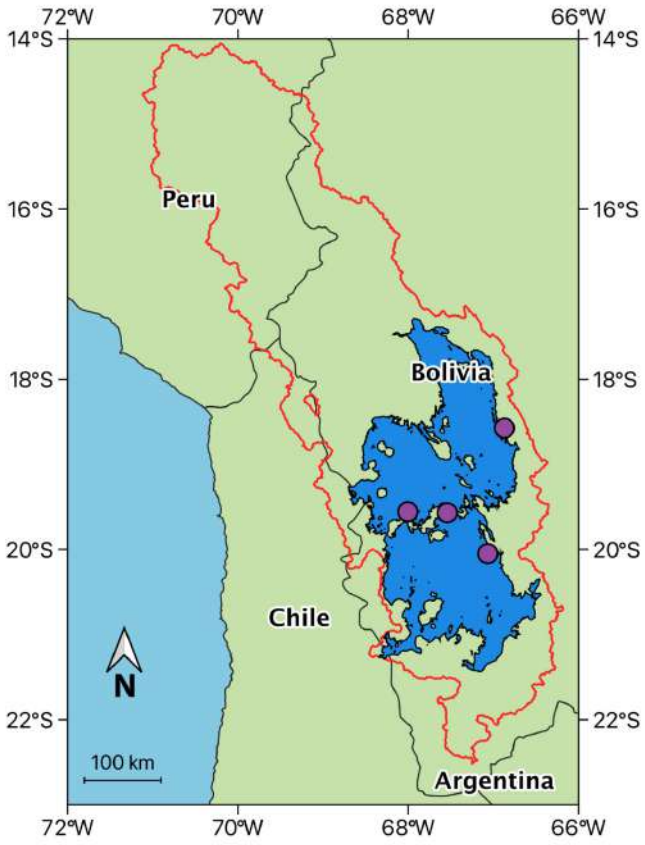


Figure 1: Representation of Lake Tauca over modern-day geography. The shape of Lake Tauca is represented by the solid blue region superimposed on a modern-day map. Lake basin is represented by a solid red line. Sample locations are represented by solid purple circles.

Geochronologic data for lake deposits indicate the Lake Tauca highstand lasted from 18 to 14.5 ka BP (the duration of HS1), with a maximum height of 3770 m from 16.5 to 15 ka (*8*). After the Tauca highstand water level dropped 70 m to about 3700 m, which corresponds to the Coipasa highstand, lasting from 12.5 to 11.9 ka BP (during the YD) (*8*).

In this work, samples from Blard *et al*. (2011) are processed and analyzed in order to reconstruct hydroclimate at Lake Tauca in the Bolivian Altiplano. The results, once compared with other findings in the field, can answer questions on the impact regional climate forcings had on hydrologic conditions during the deglacial periods of the Tauca and Coipasa highstands. Our hypothesis is that increases in precipitation (rather than decreases in evaporation) were integral to sustaining Lake Tauca, and its desiccation suggests that modern climate models can be improved by hindcasting precipitation more accurately.

MATERIALS AND METHODS

In this study, temperature, precipitation, and evaporation, which are hydroclimate factors that contribute to changes in lake levels in Lake Tauca and the surrounding South American region, were reconstructed. Carbonate clumped isotope (Δ47) thermometry, a method that looks at the degree of ordering in ^{13}C and ^{18}O bonds in carbonate minerals, was utilized to accomplish this reconstruction. The 'clumping' of heavy isotope pairs in a sample is favored at cooler temperatures, and therefore directly related to the formation temperature of the mineral (*9*, *10*). Gastropod shells, tufas, and microbialites, which were collected based on the condition that they were formed in association with constructional flat benches or shoreline notches (*8*), made of carbonate minerals, captured the Δ_{47}. The carbonate oxygen isotope composition of the samples were combined with the temperature constraints provided by Δ_{47} to estimate the oxygen isotopic composition of past lake water (*11*). This reconstructed lake water temperature is applied to a seasonal water-to-air temperature transfer function by Hren and Sheldon, which accounts for preferential carbonate formation in surface lake water during the warmer months of April to October, to estimate mean annual air temperature (MAAT; *12*). These reconstructed lake water temperatures and MAAT estimates quantitatively estimated past precipitation and evaporation rates at Lake Tauca. Precipitation and evaporation findings were corroborated by the findings of previous studies derived from paleoclimate proxies other than shoreline carbonates. Lastly, the derived estimates of hydroclimatic variables were compared to transient climate model output (TraCE-21ka; *13*), and evaluated model skill, the accuracy of the model's performance at the site, through time using the geologic benchmarks.

Sample Preparation

For tufas and microbialites, a Dremel drill was used to grind the host rock into a fine powder. The drill was set at a low speed (between 10-15 RPM) in order to prevent bond reordering caused by frictional heating. Gastropods were sonicated, dried thoroughly, then crushed into a fine powder using a mortar and pestle. Powders were reacted with a 3% hydrogen peroxide solution to remove organic material. After cleaning the samples, they were vacuumed and dried in an oven at <50°C to prevent bond reordering due to frictional heating.

Stable and Clumped Isotop Measurements

The powdered samples were weighed out to 5 mg of carbonate in silver capsules. This portioned capsule reacted with 105% phosphoric acid to produce CO_2 gas. Afterward, the gas ran through a series of purification steps. The first step was a trap containing dry ice and liquid ethanol (-76°C) which caused the water in the gas to freeze. The second step included a second trap liquid nitrogen (-196°C) which froze out the carbon dioxide in the gas. The gas moved through a silver wool column that removed sulfur into a fourth trap, a gas chromatograph which used Porapak-Q to remove additional compounds as a final step to purify the resultant CO_2. The remaining purified gas was then transferred to a mass spectrometer where a reference gas (Oztech Brand, $\delta^{13}C$ = -3.66‰ VPDB, $\delta^{18}O$ = +24.98‰ VSMOW; "$\delta^{18}O$" = ratio of ^{18}O to ^{16}O) was compared to the sample gas in order to measure for masses 44 through 49.

Sample name		Sample type	n	Age (ka)	1 s.d.	Elevation (km)	$\delta^{13}C$ (‰; VPDB)	1 s.d.	$\delta^{18}O$ (‰; VPDB)	1 s.d.	Δ_{47} (‰; I-CDES)	1. s.e	Water Temperature (°C)	1. s.e.
Tauca Highstand														
	Hua-44-1	Tufa	6	15.1	0.5	3770	4.6	0.0	-1.5	0.0	0.624	0.007	18.0	2.4
	Hua-44-2	Tufa	4	15.1	0.5	3770	4.7	0.0	-1.4	0.1	0.625	0.01	17.4	3.6
	Hua-44-3	Tufa	3	15.1	0.5	3770	4.6	0.0	-1.2	0.3	0.635	0.001	13.8	0.5
	tauca 159 ka pv white	Soft carbonate	1	15.9	0.6	3765	4.3	0.0	-2.8	0.0	0.645	0.018	14.2	4.9
	tauca 159 ka pv yellow	Hard carbonate	1	15.9	0.6	3765	4.3	0.0	-2.8	0.0	0.632	0.012	17.8	3.4
	Tauca 105	Hard carbonate	8	15.9	0.6	3765	4.3	0.1	-2.7	0.1	0.645	0.007	14.3	2.0
	Jah_12b	Bioherm	6	16.6	0.2	3670	5.1	0.0	-0.8	0.0	0.629	0.008	16.1	2.8
	Jah-14a*	Bioherm	4	14.6	0.4	3670	3.2	0.0	-1.5	0.1	0.617	0.006	20.5	2.0
Coipasa Highstand														
	B3 93-26 1a yd 19	Bioherm	1	12.9	0.2	3680	3.1	0.0	0.2	0.0	0.604	0.013	25.2	5.0
	B3 93-26 3b 20	Bioherm	1	12.9	0.2	3680	2.8	0.0	-0.7	0.0	0.632	0.014	15.1	4.9
	B3 93-26 3d 22	Bioherm	1	12.9	0.2	3680	2.7	0.0	-0.4	0.0	0.637	0.006	13.2	2.0
	B3 93-26 4 21	Bioherm	1	12.9	0.2	3680	2.5	0.0	0.3	0.0	0.632	0.006	14.9	2.1
	B-93-26-3d	Bioherm	2	12.9	0.2	3680	2.8	0.5	-0.5	0.1	0.634	0.009	14.3	3.0
	B-93-26-4	Bioherm	5	12.9	0.2	3680	2.5	0.0	0.5	0.1	0.649	0.006	9.3	1.9
	Bo-93-14-1	Bioherm	2	11.9	0.2	3670	2.6	0.8	-1.2	0.5	0.627	0.006	16.7	2.2
	Bo-93-14-2	Bioherm	2	11.9	0.2	3670	3.0	0.4	-0.6	0.7	0.611	0.003	22.7	1.1
	Bo-93-14-3	Bioherm	1	11.9	0.2	3670	2.5	0.0	0.5	0.0	0.612	0.011	22.3	4.1

Table 1: Sample information and stable and clumped isotope results. (Number of samples is represented by the column titled "n." Thousand years is represented by "ka." Standard deviation and standard error are represented by "s.d" and "s.e.," respectively. Vienna Pee Dee Belemnite is represented by "VPDB." InterCarb–Carbon Dioxide Equilibrium Scale is represented by "I-CDES".) *Sample JAH-14A was a deepwater sample from Lake Tauca and not included within temperature or hydroclimate parameter reconstruction due to disparate 14-C and U/Th ages.

Correcting Measurements

Raw isotope measurements were processed in Easotope version 20201231 (*14*). There, any non-linearity in the mass spectrometer was corrected using two different compositions of equilibrated gases at two different temperatures, 25°C and 1000°C (stochastic). Then, data were transferred into I-CDES (Intercarb Carbon Dioxide Equilibrium Scale; (*14*)) using ETH 1-3 (*15*) and internal lab standards. After, samples with anomalous Δ_{48}, Δ_{49}, which can indicate potential contamination, as well as any samples with $\delta^{13}C$, $\delta^{18}O$, or Δ_{47} values beyond 3 standard deviations from the mean were excluded.

Climate Parameter Reconstruction

To derive paleotemperatures for the samples, freshwater-based material specific calibrations were used to reconstruct water temperatures (*16*). These calibrations constrained the relationship between independent estimates of water temperature and clumped isotope-derived temperatures in modern lake carbonates. Using the temperature derived from clumped isotope analysis, lake water $\delta^{18}O$ was calculated. Mean annual air temperature was estimated using a transfer function that related April-October surface lake water temperature to account for seasonality of carbonate formation (*12*). These parameters were used within a hydrologic modeling framework that assumed isotopic steady state and water balance to derive precipitation and weighted evaporation rates following the methodology outlined in Santi *et al.* (2020) and Terrazas *et al.* (2021).

RESULTS

Lake water temperature increases between the Tauca and Coipasa highstands

In this study, Δ_{47} values were derived for 18 shoreline carbonates from Lake Tauca ranging from 3670 m to 3770 m in elevation and encompassing two significant lake and climate intervals. Δ_{47} values for the sample set ranged from 0.604-0.649‰ (I-CDES). Clumped isotope measurements (Table 1) were used to obtain estimates of water temperature (°C) and water $\delta^{18}O$ (‰; VSMOW) for both of the prominent highstands during the deglaciation at Lake Tauca. Using the Hren and Sheldon equation, an average water temperature reflecting a warm season bias of 16.2 ± 2.8°C during the Tauca Highstand and 17.1 ± 5.2 °C during the Coipasa

highstand were estimated, suggesting 0.9 ± 2.4 °C of warming between the two intervals (Figure 2). These temperatures were 2.6°C and 3.5°C degrees warmer relative to the modern mean annual water temperature of 13.6 ±1.2 °C of Lake Titicaca (*8*).

Air temperature increase between the Tauca and Coipasa highstands

Estimates of MAAT were derived using an empirical relationship between modern surface lake water temperature and air temperature for 75 lakes (*14*). Application to samples from the Tauca highstand yielded air temperatures from 5.4°C to 10.9°C and a MAAT of 8.5°C while the Coipasa highstand samples ranged from -1.0°C to 18.8°C with a MAAT of 9.2 °C. Overall, this basin experienced warmer temperatures within the transition from the Tauca Highstand (corresponding to HS1) to the Coipasa Highstand (corresponding to YD).

$\delta^{18}O$ enrichment between the Tauca and Coipasa highstands

With the concurrent additional temperature constraint provided by clumped isotope analysis and carbonate oxygen isotope measurements from the samples, lake water $\delta^{18}O$ during the Tauca Highstand was estimated to range from -1.8 to 0.7‰, with an average of -0.5 ± 1.0‰. It was found that the more modern Coipasa Highstand was slightly more enriched, with a $\delta^{18}O$ range of -0.1‰ to 3.3‰ and an average of 1.1 ± 1.3‰. This is a stark contrast to the modern measurements of incoming meteoric precipitation $\delta^{18}O$ in the region which are about -11.1‰ (*17*) and current measurements from the remainder of the desiccated lake at -3.1‰ (*18*).

Differences in precipitation and evaporation rates between the Tauca and Coipasa highstands

Using the Arnold et al. (in prep) water temperature calibration, the Hren and Sheldon (2012) transfer functions, MAAT's, and the $\delta^{18}O$ water results, local precipitation rates were calculated (*16*, *12*). The era of the Tauca Highstand (752 mm/yr) was wetter than the Coipasa Highstand (498 mm/yr), a precipitation difference of -254 mm/yr between the two highstands. Other studies showed that precipitation rates during the Tauca phase were approximately 3.3x more amplified than the modern precipitation rate of 228 mm/yr (*19*).

The Tauca highstand evaporation rate was 1713 mm/yr and the Coipasa highstand evaporation rate was 1697 mm/yr, a difference of -16 mm/yr (Table 2). Both highstands displayed an increase relative to modern evaporation rates of 1270 mm/yr (*20*). Accordingly, evaporation anomalies suggested that there was a 24.7% increase in evaporation from the Tauca to Coipasa highstands (Table 2, Figure 3).

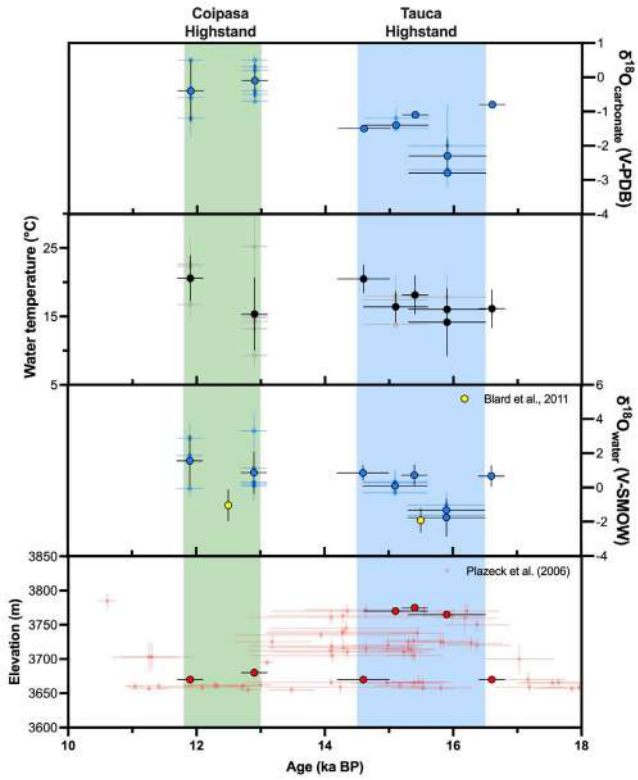


Figure 2: Carbonate $\delta^{18}O$, water temperatures, water $\delta^{18}O$, and elevation over time. The data for water temperature, $\delta^{18}O$ (carbonate and water), and elevation are plotted against radiocarbon age (ka BP) in three graphs from top to bottom, respectively. The Tauca and Coipasa highstands are differentiated by blue and green shading consistent throughout each graph. Blard et al.'s estimates for $\delta^{18}O$ are represented by solid yellow circles in order to compare the two studies (8). Placzek et al.'s estimates for elevation are represented in solid red circles in order to compare the two studies (3). (Sample size = 16) of Wnt signaling.

Comparing results with existing data

Figure 4 shows a comparison of the ratio between paleo-estimated precipitation rates and the modern mean annual precipitation rate for Lake Tauca utilizing the new data presented

Shoreline	Age (ka BP)	s.d.	MAAT (deg C)	s.d.	Temp. anomaly (deg C)	Average precipitation rate (mm/yr)	1 s.e.	Precipitation anomaly (%)	Average Evaporation Rate (mm/yr)	1 s.e.	Evaporation anomaly (%)	Average Weighted Evaporation Rate (mm/yr)	1 s.d.	Weighted evaporation anomaly (%)
Tauca	15 - 17	0.5	8.5	2.6	-5.1	752	74.4	230	1713	60.7	34.913	519.7	15.2	-59.1
Coipasa	12 - 13	0.2	9.2	6.4	-4.4	498	65.4	118	1697	96.5	33.646	205.7	11.7	-83.8

Table 2: Reconstruction of hydroclimate during Tauca and Coipasa Highstands. Anomalies were calculated using 13.6°C for modern air temperature (8), 228 mm/year for modern precipitation (19), and 1270 mm/year for weighted evaporation (20). (Sample size = 16)

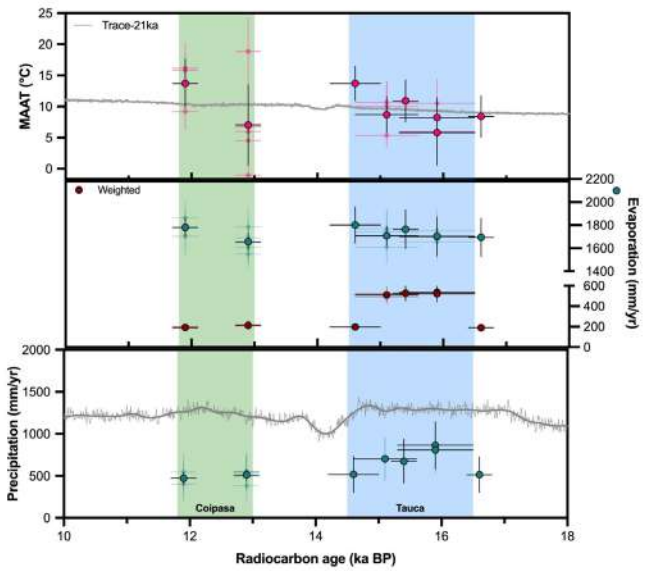


Figure 3: MAAT, precipitation and evaporation anomalies through time during both highstands. Blue shading denotes Tauca highstand while green shading denotes Coipasa highstand. MAAT is shown in the upper graph. Precipitation anomaly is shown in the middle graph, and calculated using an average annual precipitation rate of 228 mm/yr (19). Evaporation anomaly is shown in the lower graph, and calculated using an average annual pan evaporation rate of 1270 mm/yr (20). Grey line in upper and lower graphs represents findings obtained by the TraCE-21ka model. (Sample size = 16). This figure demonstrates the degree of agreement between results obtained in this study and the TraCE-21ka model. The lower skill scores for precipitation are confirmed by the distance of the plot points from the model line, which is noticeably larger than the distance between the MAAT plot points and the model line.

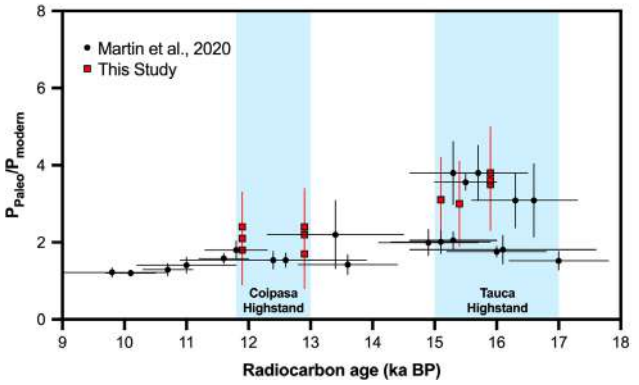


Figure 4: Comparison of derived precipitation ratio (PPaleo/PModern) for lake-based sediments (solid red squares) to estimates derived from glacier modeling (7). The findings are represented by solid red squares while Martin et al. are represented by solid black circles. (Sample size = 17). Based on this figure, the two studies are in agreement with one another, which validates what has been observed in the study, as the Martin et al. findings are derived using glacier data, not shoreline carbonates.

in this study, compared to those derived in Martin *et al.* (2020). Qualitatively, the similarity in trends between the two studies is prominent, with a noticeable dip in the ratio between P_{paleo} and P_{modern} (P_{paleo}/P_{modern}) during the Coipasa Highstand, following the peak that is reached during the Tauca Highstand. The precipitation ratio estimated by Martin *et al.* (2020), which is derived from glacial modeling, ranges between 1.2 - 3.8x modern. This range of values is within error of P_{paleo}/P_{modern} estimates from this study, which range from 3.0-3.8x modern during the Tauca highstand and 1.7-2.4x modern during the Coipasa highstand. Thus, the lake-based hydrologic modeling from this study shows a similar magnitude of precipitation forcing to those derived by moraines.

Additionally, quantitative estimates of MAAT and precipitation anomalies from this study are compared to the TraCE-21ka transient climate model, which produces continuous model output throughout the Last Glacial Maximum to modern. To do this, TraCE-21ka model output for MAAT and precipitation rates at five different intervals between 11.7 ka (Coipasa) and 17 ka (Tauca) are extracted to visually compare the data from the model (Figure 5). For the Tauca Highstand, temperatures were generally lower than the later temperatures of the Coipasa Highstand, but the region was slightly drier than the Coipasa highstand due to the slightly lower precipitation rates between the two highstands. Furthermore, a skill score (SS) is utilized to quantitatively evaluate model performance at the site (*21*, *22*). Skill scores can range from as low as -∞ to 1, with lower skill scores indicating inaccurate model performance relative to proxy estimates, while skill scores closer to 1 indicate proxy-model agreement. Overall, the models both do well at estimating past temperatures from proxy records used in this study with a skill score of 0.95 and 0.88 for the Tauca and Coipasa intervals, respectively. In contrast to the proxy-model agreement observed for MAAT, the model exhibits skill scores of 0.4 and -0.5 for evaluating precipitation during the Tauca and Coipasa highstands, respectively, suggesting that the model is not accurately capturing regional hydrologic phenomena.

DISCUSSION

Results show that evaporation decreased by 314 mm/yr and precipitation decreased by 254 mm/yr during the Tauca and Coipasa highstands. Here, how these results compare with a range of other regional paleoclimate proxies to broaden our understanding of hydroclimate drivers, including the role of ENSO on the lake's water balance, is explored. Lastly, measurements from this study are compared to a climate model (TraCE-21ka) in order to evaluate model skills using data from this study.

Comparison to Other Proxy Records

Mean annual air temperatures for both of the highstands in the Altiplano are 8.5°C and 9.2°C, for Tauca and Coipasa, respectively (Table 2). Estimates show a warming of 0.7°C between the two major highstands. These MAAT's are both cooler than modern and synchronous with the Northern Hemisphere cold periods (Heinrich Stadial 1a (HS1a), 16.5-14.5 ka BP, and YD), which are also synchronous with major advances, or standstills of paleo glaciers (*7*). This suggests that atmospheric temperatures in the Tropical Andes mimicked variability in the Antarctic. Proxy records from glacier modeling in Martin *et al.* (2020)

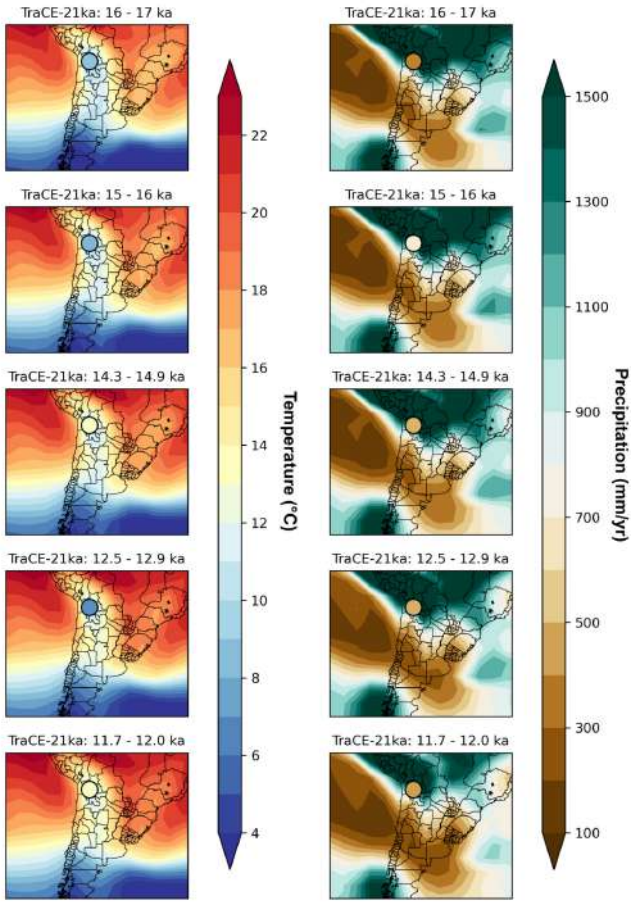


Figure 5: Climate model output from TraCE-21ka (*13*) comparison for five intervals including Tauca (17 ka) and Coipasa (12 ka) highstands. Diagrams in the left column show temperature anomalies and diagrams in the right column show modeled precipitation rates. The findings are represented by circles over the location of present-day Lake Titicaca filled in with the color represented on the scale that corresponds to the value of the results in this study. This chart allows us to visually compare the temperatures and precipitation rates found in this study with what has been hindcasted by the model.

show warming between the larger, Tauca highstand, and the desiccation to the Coipasa highstand, with an average temperature anomaly of -5°C and -3°C observed during both of the highstands, respectively (*7*). Similarly, Placzek *et al.* (2013) observed a temperature anomaly of -5.7 ± 1.1°C for the Tauca highstand alone using a coupled lake–glacier model for the Tauca lake cycle. Estimated temperature anomalies of -5.1°C and -4.4°C from this study are in agreement with these figures (Table 2).

Changes in temperature are tightly linked to changes in evaporation rates due to the Clausius-Clapeyron relationship, which demonstrates the increase in water-holding capacity of the atmosphere with rise in temperature. In this study, methodology outlined in Santi *et al.* (2020) is applied to calculate weighted evaporation, which includes evaporation from both land and free-water surfaces. Estimates from this study show suppressed weighted evaporation rates throughout the deglaciation on the Altiplano amount to a 59.1% reduction in evaporation rates during the Tauca interval, attributed to lower than modern air temperatures. Additionally, a 24.7% increase in evaporation is observed, a result of less evaporation suppression, during the Coipasa highstand relative to the Tauca interval due to the increase in MAAT's (0.7 °C, Table 2) between the highstands.

Hydrologic modeling results from this study suggest that increases in precipitation were integral to sustaining Lake Tauca at its maximum highstand. Mean annual rainfall during the Tauca interval was estimated to be 752 mm/year in this study (Table 2), a 524 mm increase relative to modern precipitation values of 228 mm/yr from Martin *et al.* (2018). Mean annual rainfall during the Coipasa highstand was estimated in this study to be 498 mm/yr, over double modern annual precipitation of 228 mm/yr (*19*).

Changes in the oxygen isotopic composition of lake water can often indicate shifts in water balance and major regional hydrologic patterns. A study by Blard *et al.* (2011) estimated paleo-$\delta^{18}O$ by assuming a formation temperature of 7.5 ± 2.5 °C, which was obtained by correcting the modern lake water temperature of Lake Titicaca (13.6 ± 1.2 °C) with a cooling of 6 ± 2 °C derived in Blard *et al.* (2009) (*2*). This lead to their estimated mean $\delta^{18}O$ for the Lake Tauca paleowater to be -2.8 ± 0.7‰ while the subsequent Lake Coipasa cycle had a mean $\delta^{18}O$ water value of -1.6 ± 0.9‰ (Blard, 2011), reflecting an enrichment of 1.2 ± 0.2‰ between the intervals (Figure 2). These values are in slight agreement with estimates of $\delta^{18}O$ derived from clumped isotope analysis in this study, with an average of -0.5 ± 1.0‰ for the Tauca Highstand and an average of 1.1 ± 1.3‰ for the Coipasa Highstand. Estimates of water $\delta^{18}O$ in this study are 2.3 ± 0.3‰ and 2.7 ± 0.4‰ more enriched for the Tauca and Coipasa highstands, respectively, than the estimates from Blard *et al.* (2011). As such, mean values from this study for each highstand end up being greater than Blard's estimations (Figure 2). Therefore, since $\delta^{18}O$ enrichment of lake waters can result from evaporation processes, both results support the notion that evaporation rates were indeed higher than seen from prior models.

Evaporation rates being higher than seen in previous models is significant as it indicates to us where models can specifically be improved in order to increase model-climate agreement.

Climate Model Comparison

Comparing results from this study with what the TraCE-21ka model predicts allows the conclusion that the Tauca Highstand was cooler yet drier than what was expected, and the Coipasa Highstand was both hotter and drier than what is expected (Figure 5) to be drawn. Modern temperatures in the Bolivian Altiplano were warmer than the temperatures estimated in this study during the two highstands, and these two cooler periods of the Tauca and Coipasa highstands were synchronous with the Northern Hemisphere cold periods (HS1a and YD), which were also synchronous with major advances or standstills of paleo glaciers (*7*). These results indicated that atmospheric temperatures in the Tropical Andes correlates with variability in the Antarctic. The warming that occurred between the Tauca highstand and the Coipasa highstand led to a 24.7% increase in evaporation during the Coipasa highstand relative to the Tauca interval, and this confirmed the hydrologic modeling results from this study, which suggested that increases in precipitation were integral to

sustaining Lake Tauca at its maximum highstand. These higher evaporation rates were also confirmed by the $\delta^{18}O$ measurements obtained in this study, which suggested a 1.6 ± 0.3‰ $\delta^{18}O$ enrichment between the Tauca and Coipasa highstands, and evaporation processes can lead to $\delta^{18}O$ enrichment of lake waters.

Sustaining Lake Tauca would have required increased precipitation, rather than a decrease in evaporation, and the fluctuations in the regional water cycle covered in this study led to its disappearance. This is in agreement with the results of the lake-based hydrologic modeling from this study, which show a smaller magnitude of precipitation forcing than those derived by glacial modeling.

Limitations

The methodology used assumes preferential carbonate formation in Lake Tauca throughout the warm season (Spring - Summer), which has been found to encompass most gastropod shell growth and biologically-mediated carbonate formation (*12*). However, if carbonate formation primarily occurs in a narrower window, MAAT's derived in this study could be slightly underestimated and be a potential source of error. MAAT's derived in this study also cover a wide range of temperatures (5.4°C to 10.7°C for the Tauca highstand and -1.0°C to 18.8°C for the Coipasa highstand) and this variability can skew MAAT results. Additionally, measurements extracted via the process outlined in the methods can lead to anomalous measurements, potentially from changing conditions in the mass spectrometer or contamination in the automated line, which naturally gives way to exclusions in the data and diminishes the size of the sample set. The large sample size requirement (roughly 5-6 mg of carbonate per analysis), prohibited additional replication of samples with limited powder (e.g. a single gastropod shell). Assumptions in modeling and additional uncertainty not solved by assumptions made could also skew precipitation and evaporation estimates as they are sensitive to model parameter uncertainty. Future work includes running additional analyses to ensure robust sample sizes for samples that can be replicated.

ACKNOWLEDGMENTS

We thank the Tripati Lab Group for their support, particularly Team Western Lakes. This work was supported by NSF and the Center for Diverse Leadership in Science, which is supported by the American Indian Studies Center, NSF, the Silicon Valley Community Foundation, the Packard Foundation, and the Sloan Foundation

REFERENCES

1. D. Palacios, *et al.*, The deglaciation of the Americas during the Last Glacial Termination. *Earth Sci*. **203**, 103113 (2020).
2. P.-H. Blard, *et al.*, Late local glacial maximum in the Central Altiplano triggered by cold and locally-wet conditions during the paleolake Tauca episode (17–15ka, Heinrich 1). *Quaternary Science Rev*. **28**, 3414–3427 (2009).
3. C. J. Placzek, J. Quade, P. J. Patchett, A 130ka reconstruction of rainfall on the Bolivian Altiplano. *Earth Planet Sc. Lett*. **363**, 97–108 (2013).
4. F. Vimeux, F. Sylvestre, M. Khodri, (2009). Past climate variability in South America and surrounding regions: from the Last Glacial Maximum to the Holocene (New York: Springer).
5. A. Terrazas, et al., Evolution of Hydroclimates in Southeast Arizona Over the Past 20,000 Years. *Undergraduate Science Journal*. **34**, (2021).
6. P. J. Bartlein, *et al.*, Pollen-based continental climate reconstructions at 6 and 21 ka: a global synthesis. *Clim. Dynam*. **37**, 775–802 (2011).
7. L. C. P. Martin, *et al.*, Antarctic-like temperature variations in the Tropical Andes recorded by glaciers and lakes during the last deglaciation. *Quaternary Science Rev*. **247**, 106542 (2020).
8. P.-H. Blard, *et al.*, Lake highstands on the Altiplano (Tropical Andes) contemporaneous with Heinrich 1 and the Younger Dryas: new insights from 14C, U–Th dating and δ18O of carbonates. *Quaternary Science Rev*. **30**, 3973–3989 (2011).
9. J. M. Eiler, "Clumped-isotope" geochemistry—The study of naturally-occurring, multiply-substituted isotopologues. *Earth Planet Sci. Lett*. **262**, 309–327 (2007).
10. E. A. Schauble, P. Ghosh, J. M. Eiler, Preferential formation of 13C–18O bonds in carbonate minerals, estimated using first-principles lattice dynamics. Geochimica et Cosmochimica Acta. **70**, 2510–2529 (2006).
11. L. M. Santi, et al., Clumped isotope constraints on changes in latest Pleistocene hydroclimate in the northwestern Great Basin: Lake Surprise, California. GSA Bulletin. **132**, 2669–2683 (2020).
12. M. T. Hren, N. D. Sheldon, Temporal variations in lake water temperature: Paleoenvironmental implications of lake carbonate δ18O and temperature records. *Earth Planet Sc. Lett*. **337–338**, 77–84 (2012).
13. Z. Liu, *et al.*, Transient Simulation of Last Deglaciation with a New Mechanism for Bølling-Allerød Warming. *Science*. **325**, 310–314 (2009).
14. C. M. John, D. Bowen, Community software for challenging isotope analysis: First applications of 'Easotope' to clumped isotopes: Community software for challenging isotope analysis. *Rapid Commun. Mass*. Sp. **30**, 2285–2300 (2016).
15. S. M. Bernasconi, *et al.*, InterCarb: A Community Effort to Improve Interlaboratory Standardization of the Carbonate Clumped Isotope Thermometer Using Carbonate Standards. *Geochem. Geophy. Geosy*. **22** (2021), doi:10.1029/2020GC009588.
16. A. J. Arnold, *et al.*, AGU Fall Meeting, Washington, D.C., Dec. 10-14, 2018. PP13F-1396.
17. G. J. Bowen, The Online Isotopes in Precipitation Calculator, version 3.1. (2022) (http://www.waterisotopes.org)
18. R. Gonfiantini, *et al.*, "Hydrochemical and isotope study of Lake Titicaca" in Study of Environmental Change using Isotope Techniques. (2001)
19. L. C. P. Martin, *et al.*, Lake Tauca highstand (Heinrich Stadial 1a) driven by a southward shift of the Bolivian high. *Sci. Adv*. **4**, eaar2514 (2018).
20. T. A. Blodgett, B. L. Isacks, J. D. Lenters, Constraints on the origin of paleolake expansions in the Central Andes. *Earth Interact*. **1**, 1–28 (1997).
21. J. C. Hargreaves, J. D. Annan, R. Ohgaito, A. Paul, A. Abe-Ouchi, Skill and reliability of climate model ensembles at the Last Glacial Maximum and mid-Holocene. *Clim. Past*. **9**, 811–823 (2013).
22. J. M. Lora, Components and mechanisms of hydrologic cycle changes over North America at the Last Glacial Maximum. *J. Climate*. **31**, 7035–7051 (2018).

Accelerated wound healing through downregulation of Neuronal Pas Domain 2

Adam Clements[1], Reza Jarrahy, PhD[2,3], Akishige Hokugo, PhD[2,3]

[1]Department of Bioengineering, [2]Division of Plastic and Reconstructive Surgery, Department of Surgery, [3]David Geffen School of Medicine at UCLA

ABSTRACT

The core circadian clock gene, Neuronal PAS Domain 2 (Npas2), has been shown to significantly accelerate dermal wound healing while mitigating the formation of hypertrophic scars (HTS). Npas2 is also responsible for controlling the active phase of wound healing in which tissues produce collagen to facilitate wound closure and prevent infection; however, these mechanisms may result in HTS formation from excessive collagen. To combat this, five different hit compounds that can downregulate Npas2 activity were examined. In this study, therapeutic downregulation of the Npas2 gene was hypothesized to have three effects: 1) downregulate collagen synthesis genes, 2) decrease overall collagen biosynthesis, and 3) accelerate dermal wound healing without the formation of HTS. To test the effects of each hit compound (named Dwn1, 2, 3, 4 and 5), primary human fibroblasts were treated with either 0, 0.1, 1, or 10 μM of a single hit compound. Control groups were untreated with the negative control grown in media lacking ascorbic acid (A.A.) supplementation to serve as an indicator of stable collagen formation. Comparison of control groups allowed for the examination of the role of A.A. in the formation of these collagen structures. Groups were then observed for total collagen deposition by Picrosirius Red staining and activity of collagen type I and type III genes through quantitative RT-PCR. Dwn1 was used in cell migration assays to analyze how drug treatment alters cell infiltration into the wound site. Results from this study indicated that both Dwn1 and Dwn5 may serve as possible therapeutic agents for future studies related to dermal wound healing. This research has implications in accelerating surgical recovery which could improve clinical outcomes by reducing hospital stays and easing the financial burden faced by surgical patients due to extended treatment.

INTRODUCTION

Circadian clock genes maintain cellular metabolism in response to environmental stimuli (1), dictate the 24-hour sleep/wake cycle that is controlled by the suprachiasmatic nuclei (SCN), and determine the ability of tissues to maintain metabolic cycles (1). For the purposes of this study, the Neuronal Pas Domain 2 (Npas2) gene was studied in isolation to better understand its role in wound healing, thereby elucidating the role of circadian biology as a whole within the wound healing process. Specifically, the role of Npas2 was studied to determine the way in which downregulation of circadian genes is able to prevent the formation of hypertrophic scars (HTS). These kinds of scars result from trauma, surgery, inflammation, and burns due to scars crossing over the skin's natural creases, causing the misalignment of collagen proteins (2). The formed scar is often raised and discolored, exhibiting abnormally high levels of collagen types I and II (C1, C3) in different ratios to normal skin (3). Regulation of these collagen types presents an opportunity to reduce the prevalence of HTS following surgical operations through regulation of the Npas2 gene. Therefore, this study hypothesizes that the downregulation of Npas2 will have three effects: 1) downregulate collagen synthesis genes, 2) decrease overall collagen deposition, and 3) accelerate dermal wound healing without HTS. The objective of this study is to determine the extent to which the five hit compounds regulate collagen gene expression and overall collagen synthesis.

This study utilized the Npas2 downregulation of hit compounds to investigate the impact of this gene on wound healing, looking at gross collagen deposition following therapeutic treatment and the impact of each hit compound on the collagen synthesis genes themselves. Following this, one particular compound (Dwn1) was selected to perform a 48-hour experiment, showing the overall circadian gene activation and cell migration during this time period. In total, two compounds (Dwn1 and Dwn5) were found to be promising options for future wound healing studies relating to the impact of mitigated circadian activation during the wound healing cycle.

This study has implications across multiple fields of study, primarily focusing on circadian biology and wound healing. By influencing the way in which surgical wounds heal, accelerated wound healing is able to provide clinical benefits while also mitigating cost. Ranging from physical benefits, including a reduced chance of infection and a greater degree of biofunction

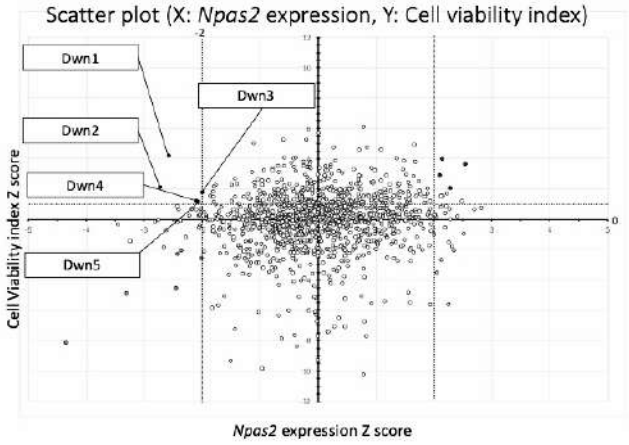


Figure 1. Identification of 5 Hit Compounds through Npas2 Reporter System and Cell Viability Testing. Dermal fibroblasts were with an engineered LacZ reporter gene within the *Npas2* allele. Npas2-LacZ reporter system and cell viability tests were conducted, plotting the Z-score for each via dotted lines. High absolute values of negative Z-score for *Npas2* were used as indicators of significant gene downregulation (x-axis), while high cell viability Z-scores were used to indicate that fibroblasts undergoing treatment remained viable through the course of the experiment (y-axis). The noted compounds were shown to reduce *Npas2* expression and were selected for having an *Npas2* expression Z-score of less than -2 and a cell viability index Z-score of greater than 1.

following the wound healing process, to the psycho-social benefits of preventing the formation of unwanted facial scarring, improvements to this field could vastly improve the individual patient experience. These treatments have the added benefit of lowering overall surgical costs by reducing the amount of time spent in the hospital and the need for additional procedures following unsuccessful healing. By implementing these hit compounds into clinical practice, surgical wounds may be able to heal faster and produce substantial benefits to the patient during the recovery process following surgical procedures.

MATERIALS AND METHODS

Statement of Animal Use

All protocols for animal experiments were approved by the University of California Los Angeles (UCLA) Animal Research Committee (ARC# 2003–009) and followed the Public Health Service Policy for the Humane Care and Use of Laboratory Animals and the UCLA Animal Care and Use guidelines. The animals were fed a regular rodent diet and provided water ad libitum. They were maintained in regular housing conditions with 12-hour light/dark cycles in the Division of Laboratory Animal Medicine at UCLA.

Hit compound identification

At the Molecular Screening Shared Resource (MSSR) at UCLA, a drug library of 1,120 FDA-approved compounds was screened with two different assays to identify hit compounds with wound healing properties. Hit compounds involved in the modulation of murine dermal fibroblast Npas2 expression were identified using high-throughput screening of gene activity levels. Dermal fibroblasts were isolated from mice engineered to carry the LacZ reporter gene in the Npas2 allele. LacZ reporter gene activity has previously been shown to accurately correlate with endogenous Npas2 expression (4). The cells were cultured in growth medium containing Dulbecco's Modified Eagle's medium (DMEM; 11995065, Life Technologies Corp.) with 10 % fetal bovine serum (FBS; 1600004, Life Technologies Corp.) and 1% penicillin/streptomycin (PS; 15140122, Life Technologies Corp.). Using 384-well plates (781906, Greiner Bio-One) and a pin tool (Biomek FX, Beckman Coulter), each well was filled with 25 µl non-phenol red DMEM (31053036, Life Technologies Corp.), which contained 10% FBS and 1% PS, and 50 nL of FDA-approved compounds, giving a final concentration of 1µM. Cells were added to each well (1,500 cells per 25µl) and incubated at room temperature for 1 hour, followed by a 48 hour incubation at 37ºC and 5% CO_2. To measure Npas2-LacZ expression, β-galactosidase activity was measured using a Beta-Glo Assay System (E4720, Promega). The Npas2-LacZ expression data were uploaded to an online data analysis tool (CDD Vault, Collaborative Drug Discovery Inc), on which data were normalized and the Z-factor was calculated. Separately, the same compound library was screened for human dermal fibroblast migration. A commercially available human dermal fibroblast cell line (CCD-1122Sk, ATCC; 3,000 cells per 25µL) was applied on OrisTM Pro Cell Migration Assay 384-well plate (PRO384CMA1, Platypus Technologies), which has a water-soluble biocompatible gel that creates a center cell-free detection zone for cell migration in each well. After plating the cells, the plates were centrifuged at 200 rotations per minute (RPM) for 5 min. After a 1-hour incubation at room temperature for cell attachment, the compounds were added using a 250 nL pin tool and incubated at 37ºC in the CO2 incubator. After 48 hours of incubation, 25 µl of staining solution (Calcein-AM and Hoechst, Life Technologies Corp.) was added to each well. After another period of centrifugation at 200xg for 5 min, the plates were incubated for 20 min at room temperature, and each well was then imaged by the Micro Confocal High-Content Imaging System (ImageXpress, Molecular Devices). The cells that migrated into the detection zone were counted using a customized computer program (CDD Vaultâ, Collaborative Drug Discovery, Burlingame, CA) and the Z-factor was calculated.

Dermal Fibroblast Cell Culture

Primary human fibroblasts (hDFa) isolated from a healthy 20-year-old female were obtained from ATCC. Cells were cultured in growth medium including modified DMEM treated with 10% FBS and 100 units of penicillin per 0.1 mg/mL streptomycin at 37°C and 5% CO2 in a humidified incubator. In addition, growth media was supplemented with A.A. (50 µg/mL; Sigma-Aldrich Corp.). All experimental groups were cultured in growth media supplemented with A.A. as this compound has been shown to be essential for the development of a normal, strong, and mature collagen network as well as maintaining its optimal collagenic density (5). Medium changes were conducted every 3 days.

Name	Action
Dwn1	Antibiotic (pesticide that blocks mETC complex III, causing cell death)
Dwn2	Analgesic/Anti-inflammatory (used to treat rheumatoid arthritis by inhibiting phospholipase A2 and COX-2)
Dwn3	Antagonistic (Antipsychotic used to treat schizophrenia)
Dwn4	CNS Stimulant
Dwn5	Antagonistic (Used to treat hypertension, psychiatric disorders, and tardive dyskinesia)

Table 1. Mechanism of Action and Current Use for Each Hit Compound Identified. The general mechanisms of action and the clinical conditions commonly treated using the drug compounds Dwn1-5.

Collagen expression by Picrosirius red staining

The data collected for this experiment relied on the principles outlined by the Beer-Lambert Law, which states that increases in concentration will result in linear increases in absorbance values. For this experiment, the absorbance was set to 550 nm to calculate the collagen concentration stained by Picrosirius Red. hDFa cells were seeded in a 24-well plate in the growth medium with A.A. (50 μg/mL) and various concentrations (0, 0.1, 1, and 10 μM) of the hit compounds. The negative control was placed in identical conditions, except that its media did not include A.A. supplementation. On day 7 and 14, the cells were fixed with 10% neutral buffered formalin and stained with Picrosirius Red (PolyScience, Niles, IL) for gross collagen visualization. For a quantitative analysis, the stained dye was then eluted in 0.1 N sodium hydroxide from each well, and a plate reader (SYHNERGY H1 plate reader) was used to determine the absorbance at a wavelength of 550 nm. Findings were compared by one-way ANOVA and Dunnett's Test.

Gene expression

RNA samples were extracted from cultured hDFa on experimental day 7 using the RNeasy kit (Qiagen, Valencia, CA), and were then analyzed for quality and concentration using NanoDrop (Thermo Fisher Scientific Inc.). cDNA synthesis was carried out using the High-Capacity RNA-to-cDNATM kit (ThermoFisher Scientific Inc.) following the manufacturer's protocol. Quantitative real-time PCR (qPCR) was carried out using TaqMan primer/probe sets (Thermo Fisher Scientific Inc., Waltham, MA), as per manufacturer instructions, for genes C1 (Col1a1: Mm0080166_g1, Col1a2: Mm00483888_m1) and C3 (Col3a1: Mm00802300_m1). Statistical analysis was performed using one-way ANOVA and Dunnett's tests, using the positive control as the point of comparison for determination of statistical differences between groups within each experiment.

Npas2 gene expression with Dwn1 treatment

Primary dermal fibroblasts derived from WT mice were harvested and cultured as previously described. Fibroblast cultures were treated with 10 μM Dwn1. After synchronization using 10 nM dexamethasone, all RNA from the fibroblasts had cultured while 10 μM Dwn1 was extracted every 6 hours from hour 12 to 42 (RNeasy® Plus Mini kit, Quiagen), followed by cDNA synthesis (SuperScript® VILOTM cDNA Synthesis Kit, Thermo Fisher Scientific). Taqman-based qRT-PCR was performed using a primer/probe mix, Npas2 (Mm01239312_m1, Thermo Fisher Scientific), and an endogenous control mix comprised of mouse Gapdh (4352339E, Thermo Fisher Scientific). Statistical analysis was performed by two-way ANOVA.

RESULTS

Identification of hit compounds

Each compound's Npas2 downregulation Z-score was correlated to their respective cell migration Z-score following the drug screening of the different compounds available through the MSSR. A high absolute value of the Npas2 Z-score was used as an indicator of significant gene downregulation (Figure 1, x-axis), while high cell viability Z-scores were used to indicate that fibroblasts being treated remained viable through the course of the experiment (Figure 1, y-axis). In order to select compounds, those with a Npas2 gene activity Z-score greater than -2 and those with a cell viability Z-score less than 1 were eliminated from consideration. A -2 Z-score value indicates an Npas2 downregulation within the 98th percentile of the all compounds tested, and a Z-score of 1 indicates a cell viability in the 84th percentile. Together, these selection criteria were able to weed out all but the five selected hit compounds. The five compounds were used as possible therapeutic agents and assigned identifying names in order of decreasing cell viability index Z-score values. The mechanisms of action for these compounds were then identified and recorded in Table 1. Although some hit compounds share similar overall mechanisms of action– for example, Dwn3 and Dwn5 are both antagonistic– the specific cellular signaling pathways employed by each compound differ, thereby causing different levels of effect to the Npas2, and thus, different levels of collagen deposition in the wound site.

Overall Collagen Biosynthesis

The indication of each compound's ability to serve as a treatment for

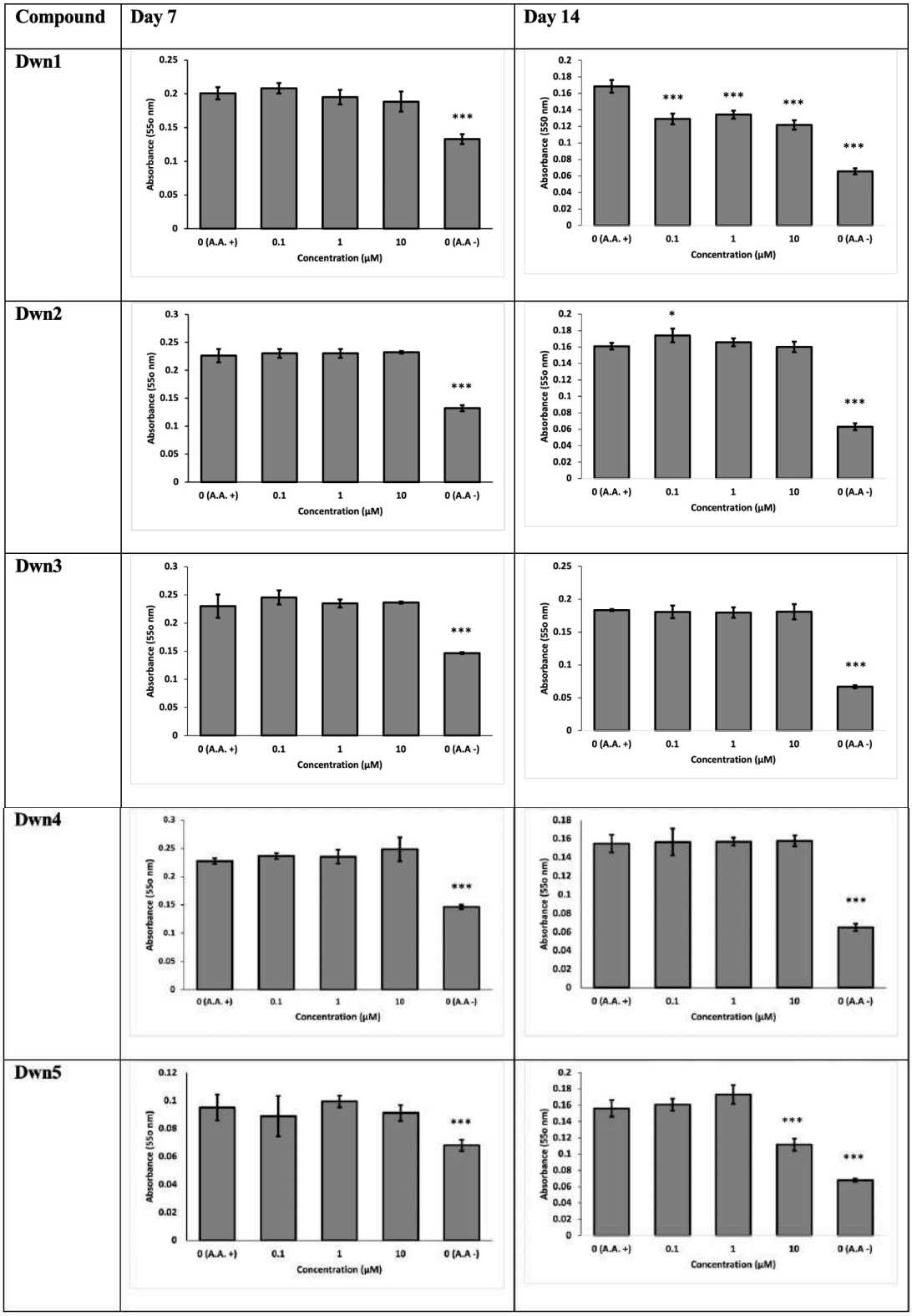


Figure 2. Collagen Expression on Day 7 and Day 14 After Hit Compound Treatment. Absorbance values were calculated from spectrophotometer data after Picrosirius Red staining of primary human dermal fibroblasts (hDFa). Absorbance values were determined from collagen concentrations, with higher levels of total collagen expression eliciting higher absorbance values due to the red staining. Images were taken on day 7 (D7) and day 14 (D14) (n=4). The positive control was the 0 μM group cultured with ascorbic acid (A.A.) supplementation in the growth media (50 μg/mL). The negative control was the 0 μM group cultured without A.A. supplementation in the media. Dwn1 (D7 and D14) and Dwn5 (D14) can be seen to significantly decrease gross collagen deposition. Results show experimental group average ± STD. Analysis consisted of Dunnet's Test and one-way ANOVA (*P<0.05, **P<0.01, ***P<0.001; significant difference determined in comparison to positive control; bars are representative of standard error).

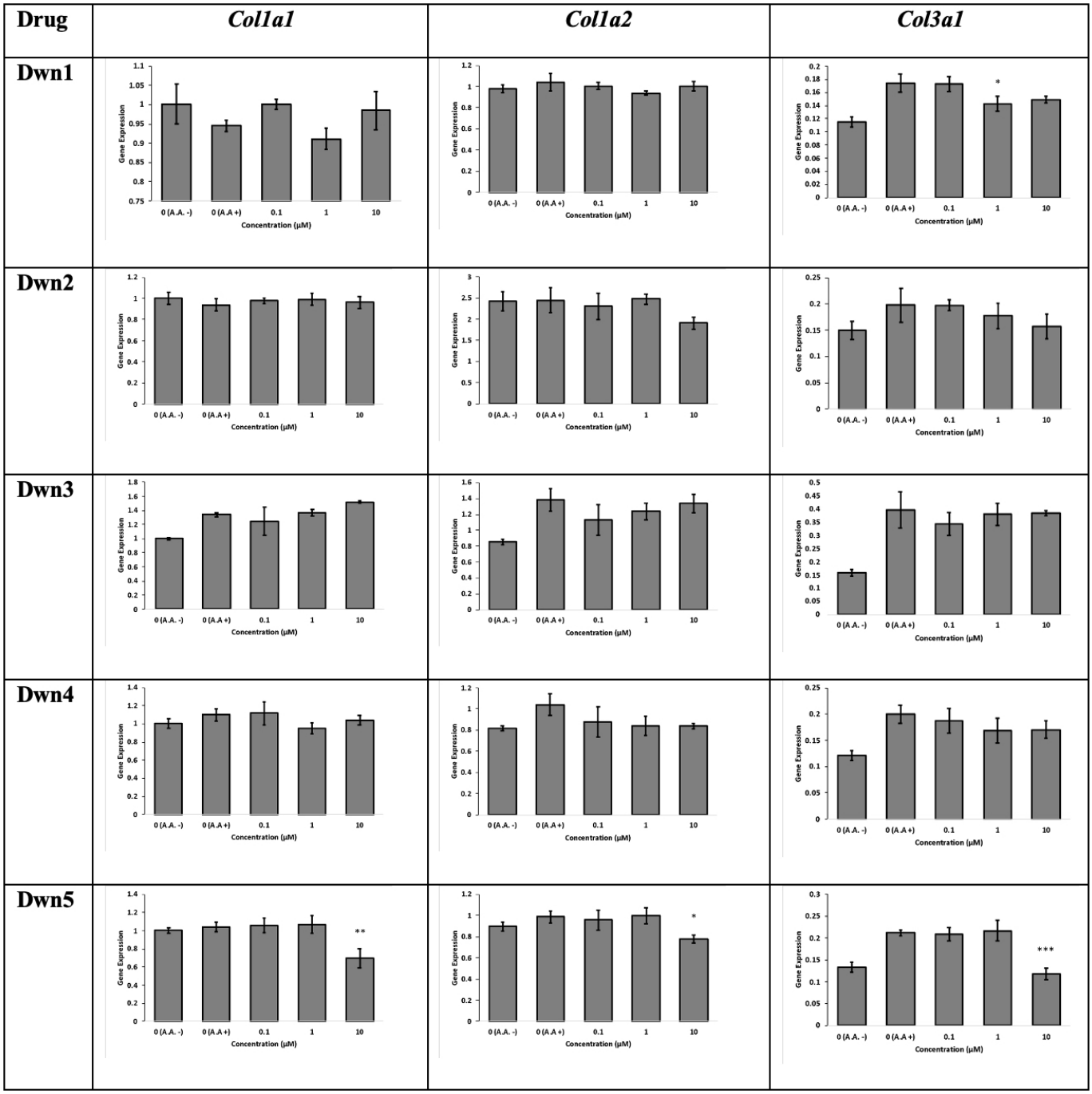


Figure 3. Collagen Synthesis Gene Expression Downregulation for Dwn 1 and 5 at Various Concentrations. Gene expression of C1 (*Col1a1, Col1a2*) and C3 (*Col3a1*) relative to the positive control group (0 μM (A.A. +). The gene expression for each of the experimental groups was calculated using quantitative PCR (qPCR) on experimental day 7. The graphs represent the average of each test case (n=4) ± STD. Data recorded pertaining to negative controls were not included in significance calculations, but were used as a reference to determine how ascorbic acid (A.A.) deficient cells respond to hit compound treatment. Significant downregulation of genes can be seen for Dwn 1 and Dwn 5 treated groups. The gene expression values analysis consisted of Dunnet's Test and one-way ANOVA (*P<0.05, **P<0.01, ***P<0.001).

the prevention of HTS formation was determined from the overall collagen biosynthesis. The results from Figure 2 indicated that although there was no difference between the positive control and any experimental group on day 7, there were notable trends for some groups. Dwn1, Dwn3, and Dwn4 all showed slight increases in the 0.1 μM treatment groups. Although these findings suggested an increase in collagen deposition, the change was insignificant and may have resulted from experimental error or an unintended side effect of drug treatment (p-value > 0.05). The cause of these trends were not explored further in the present study. The significant difference noted between the positive and negative control groups can be attributed to the lack of A.A. supplementation in the negative control media. A.A. has been shown to be important for the development of strong collagen fibers as the collagen that was formed in media lacking A.A. were more prone to damage and instability. For this reason, the overall collagen biosynthesis for the negative control was significantly lower for all groups. The control group lacking A.A. supplementation was included to accurately identify the produced proteins as collagen since a significant decrease was expected when compared to the experimental groups due to the protein instability. Values similar to those of the positive control indicated the presence of structural proteins other than collagen, thereby indicating a different mechanism of action for the chosen hit compounds.

As shown in Figure 2 data for day 14, Dwn1 (0.1, 1, 10 μM) and Dwn5 (10 μM) both displayed increased collagen deposition compared to the A.A. positive control group. Dwn2 (0.1 μM, $p<0.05$) showed an increase in overall collagen deposition compared to A.A. positive control, despite higher drug concentrations of the therapeutic agents (1 μM and 10 μM) resulting in no statistical difference. None of the concentrations of Dwn4 and Dwn3 showed a change in collagen deposition throughout the duration of the experiment.

On day 7 and day 14, some hit compound concentrations (0.1 and 1 μM) showed insignificantly higher absorbance values in comparison to the positive control group (p-value > 0.05). This trend could be seen with the day 7 results from all five hit compounds, as well as Dwn2, Dwn4, and Dwn5 on day 14. While these results were not statistically significant, they suggest that low doses of Npas2 downregulation agents may result in a slight increase of overall collagen production in the local microenvironment. Although the method for calculating concentration relied on the Beer-Lambert Law, the impact of the hit molecules on overall collagen deposition did not follow the same trend as there was not a linear proportional relationship demonstrated between the lower and higher concentration groups.

Collagen Synthesis gene expression

Relative gene expression of C1 (Col1a1, Col1a2) and C3 (Col3a1) was measured on day 7 to determine the effect of each hit compound on the specific genes related to collagen deposition. As shown in Figure 3, each compound was used to treat different groups, yet only some resulted in significant changes between the positive control and experimental groups. Dwn1 (1 μM) showed a decrease in Col3a1 expression. Dwn2, Dwn3, and Dwn4 all showed no significant change in collagen gene expression levels (p-value > 0.05). Dwn5 (10 μM) showed significant decreases in gene expression for all three genes being tested with varying levels of significance (p-value < 0.00001). The remainder of the treatment groups showed no significant difference when compared to the positive control (p-value > 0.05).

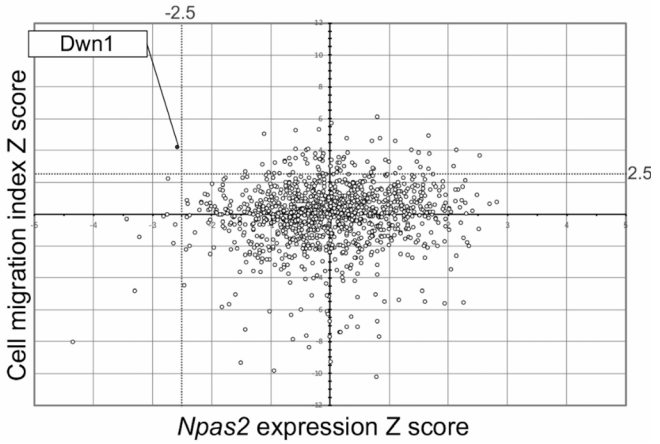


Figure 4. Impact of Dwn1 on Cell Migration and Npas2 Activity. The graph shows Dwn1 results from preliminary testing. Dotted lines indicate Z-scores for both the cell viability index (Z-score = 2.5) and the *Npas2* expression (Z-score = -2.5). Dwn1 is shown to have the highest cell viability of accepted hit compounds (Z-score = 4) while maintaining an Npas2 downregulation within the desired threshold.

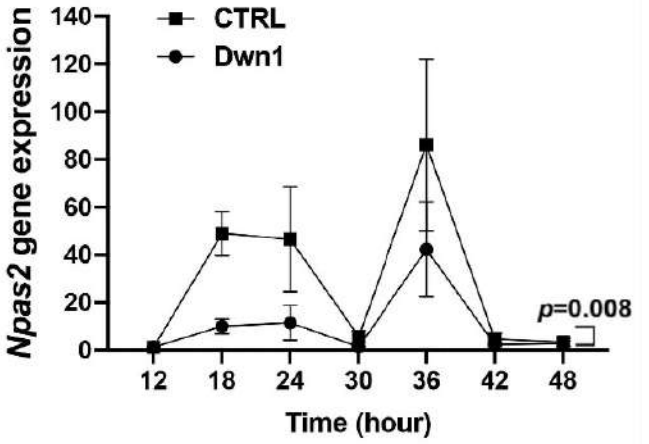


Figure 5. 48 hour Dwn1 Treatment Downregulates Npas2 Expression While Maintaining Adequate Circadian Cycle. The graph shows the Npas2 expression levels during 48 hours of treatment by Dwn1 (10 μM) with treatment groups exhibiting significant downregulation of Npas2 gene activity. Normal cyclic circadian rhythm changes can be seen by the sinusoidal nature of control and treatment groups throughout 48 hours. Analysis consisted of Dunnet's Test and one-way ANOVA. Graph represents the average ± SE. The p-value between the groups was also calculated (p=0.008).

48-hour Npas2 expression

Dwn1 was used to study cell migration and Npas2 gene expression over 48 hours because the results from HTS showed that it yielded the second lowest Npas2 downregulation Z-score (-2.601) while also maintaining the highest overall cell viability index Z-score (4.06; Figure 4) Although Dwn5 was also shown to result in the significant downregulation of the gross collagen production and

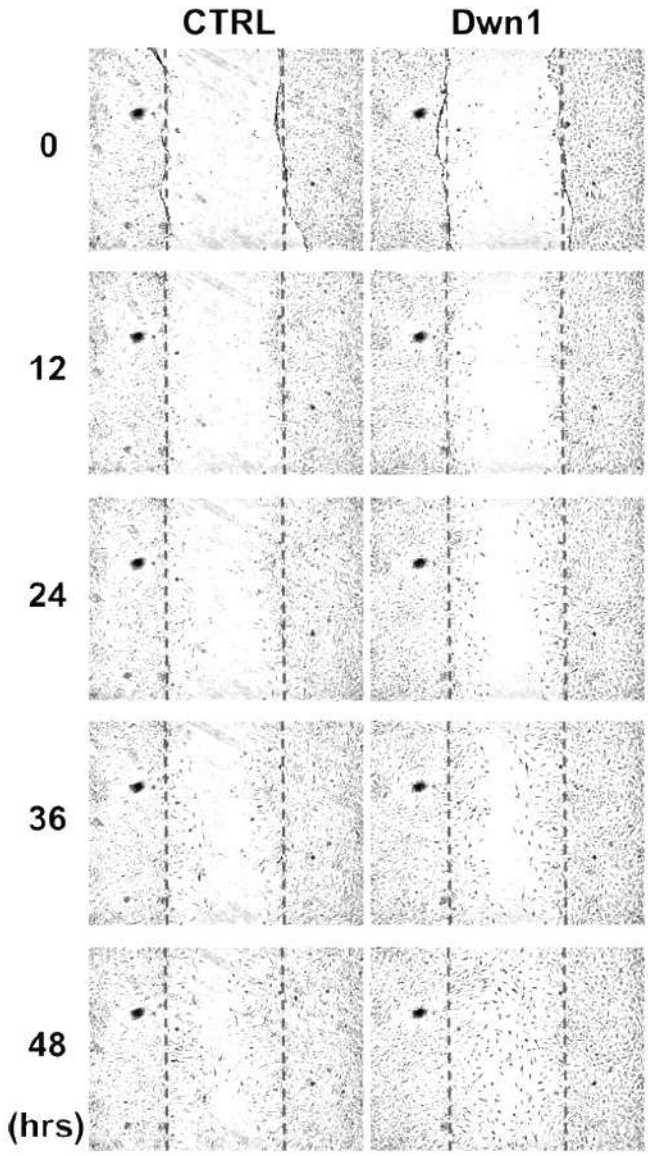


Figure 6. Dwn1 Treated Cell Migration Assay over 48 Hours via Scratch Test. Images show the cell migration of dermal fibroblasts at 12 hour intervals over the course of 48 hours. The number of migratory cells was counted (n=4) to determine the rate at which cells enter simulated wound sites. Dwn1 treated groups can be seen to have a higher overall amount of migrated cells beyond the borders of the initial scratch test.

collagen synthesis gene activity, Dwn1 showed a decrease in gross collagen deposition for all experimental concentrations as well as a decrease in all collagen synthesis genes. While Dwn5 was shown to produce significant downregulation of collagen synthesis genes at 10 μM, the values for the 0.1 and 1 μM gene expression levels were slightly higher than the positive control despite being statistically insignificant ($p>0.05$). Dwn5 showed a statistical decrease in gross collagen deposition for only 10 μM while the absorbance values for the 0.1 and 1 μM treatment groups showed slight increases. In summary, Dwn1 and Dwn5 were shown to be promising therapeutic agents; however, Dwn1 was used for the cell migration assay and 48-hour treatment as prior experiments proved Dwn1 to be better suited for further experimentation in this area.

In addition to altering the way in which collagen synthesis genes are affected after treatment with hit compounds, Figure 5 shows the ability of Dwn1 to impact the expression of the Npas2 gene over the course of 48 hours. Since many dressing applications occur during postoperative care, the hit compound used must be able to produce a sustained effect throughout the course of treatment in order to elicit a greater response. The experimental group showed a similar pattern of cyclic activity characteristic of normal circadian rhythm genes, as demonstrated by Figure 6. Figure 7 shows how the use of Dwn1 exhibited a significantly higher level of migratory cells throughout the 48 hour treatment when compared to the control group (p-value = 0.0051).

DISCUSSION

All five hit compounds were correlated with a downregulation in Npas2, as shown in Figure 1. In this study, a novel approach has been put forward for the downregulation of the core circadian clock gene Npas2 in order to reduce the amount of collagen expressed during dermal wound healing. The hypothesis was that the downregulation of collagen synthesis genes due to the different hit compounds will accelerate dermal wound healing and reduce the formation of HTS. Through in vitro experiments with all five hit compounds, Dwn1 and Dwn5 have been shown to be possible therapeutic agents for future treatments of incisional wound healing. The selected hit compounds were responsible for downregulation of Npas2 and maintained a high cell viability, yet all have different mechanisms of action. Dwn1 works by inhibiting cellular respiration through inhibiting Complex III (QH2 cytochrome c reductase) in oxidative phosphorylation, thus preventing the successive reduction of components following this pathway (2). Dwn2 follows a different pathway within the cell where it acts as a cyclooxygenase-2 inhibitor to reduce inflammatory pain in conditions such as rheumatoid arthritis, osteoarthritis, and physical trauma (6). The exact cellular pathway taken by this molecule to interact with the circadian genes is not currently understood. This same phenomenon exists for Dwn3, as this compound was developed first for clinical treatment of schizophrenia management and is now used for the treatment of impulsive aggression in children with attention-deficit/hyperactivity disorder (7). While the cellular pathway for this molecule may be unclear, its continued use in clinical research continues to be funded by the US FDA for animal testing, focusing mainly on repeated-dose toxicology studies (7). Dwn4 is a known central nervous system stimulant, yet very little information is known about how the drug interacts with human cells as this compound is currently only available for research purposes (8). Finally, Dwn5 is currently used as a hypertension medication as it is an adrenergic uptake inhibitor (4). Like the other compounds, the exact mechanism of how Dwn5 interacts with specific circadian genes remains unclear. Notably, the drug screening did show Npas2 downregulation for each of these compounds, making their role in the circadian cycle an avenue for future research.

The circadian locomotor output cycles kaput, Npas2, and aryl hydrocarbon receptor nuclear translocator-like protein 1 (Arntl,

Bmal1) genes provide transcription factors to express the period (Per2) and cryptochrome (Cry1 and Cry2) genes (9,10). Following this, Per2 and Cry1 control the 24 hour sleep/wake cycle and the wound healing cycle. The absence of the 24-hour circadian control causes fibroblast and keratinocyte hyperproliferation and under expressed collagen (11).

Npas2 is solely expressed in murine dermal fibroblasts and is augmented in peripheral tissues, showing how Npas2 is important for regulation of peripheral tissue circadian cycles (1,12). These characteristics allow for Npas2 to be studied in isolation, thereby elucidating the mechanism of action Npas2 exhibits in a wound healing model.

Mice lacking the Npas2 gene (Npas2 -/-) exhibit accelerated dermal wound healing mechanisms and increased cellular migration and contraction in vitro (1). Due to Npas2 being deleted from subjects with functioning Clock genes, wounds were evaluated in peripheral tissues with an absent circadian control at a cellular level but with an in-tact circadian control of the sleep/wake cycle. Prior findings show that injuries occurring in the active phase of cellular metabolism have a significantly faster recovery rate compared to the dormant phase (13,14). Npas2 suppression allows for a 24-hour active cellular metabolism, thereby accelerating wound healing (15). In fibroblasts, Npas2 regulates C1 and C3 in response to damage, as these proteins comprise the skin extracellular matrix (ECM) and are responsible for scar formation (2). Denoted as being trimeric ECM proteins generated from expression of Col1a1, Col1a2, and Col3a1 (16), C1 and C3 are the primary components in skin ECM due to their ability to form thicker fibers in comparison to other collagen types (6). These proteins change their orientation following injury (7), as HTS show bundles with a flatter and more wavy pattern in relation to the epithelial surface rather than parallel (14). Normally C1 and C3 exhibit a defined ratio (6) but HTS alter this ratio between collagen types (2).

This study used two groups of controls, a negative control due to the absence of A.A. and a positive control due to the presence of A.A. This supplementation is an important factor for the stabilization of existing collagen networks and was manipulated to indicate if the structural proteins found were truly collagen. For the purposes of this experiment, low levels of collagen expressed in the negative control group with respect to the other groups tested indicated the presence of stably formed collagen in all groups containing A.A. supplementation since the negative control group was expected to contain lower collagen expression levels due to the collagen fibers being unstable. The control group lacking A.A. supplementation was included to accurately identify the produced proteins as collagen, as the low value is indicative of collagen instability. Values similar to those of the positive control indicated the presence of structural proteins other than collagen, thereby indicating a different mechanism of action for the chosen hit compounds.

Dwn1 and Dwn5 both displayed a significant suppression of collagen expression overall in comparison to the positive control group. Dwn5 was responsible for causing a biphasic effect when used to treat the fibroblasts as treatment with 1 μM resulted in an increase of overall expression while treatment with 10 μM resulted in a decrease overall. The cause of such findings was not investigated in the current study. Dwn5 also showed significant downregulation of all genes relating to collagen synthesis for both C1 and C3. In addition, Dwn1 showed downregulation of C3 expression with an upward trend for C1 expression as shown by Col1a2 specifically. The paradoxical increase in gross collagen deposition for low doses of therapeutic agents may be due to the lower hit compound concentrations being unable to elicit a significant response in Npas2 gene activity, thereby allowing the native tissues to produce normal levels of collagen in response to the inflicted injury. Furthermore, this experiment used concentration of hit compounds (identifiable on the x-axis) to diminish the amount of collagen formed by hDFa cells, which are then stained to allow for spectrophotometer data collection relating to the amount of collagen present, thus the experiment is expected to result in an inversely proportional trend between hit compound concentration and absorbance. This explains why Dwn1 showed similar levels of collagen deposition for multiple concentrations on day 14 since the different amounts of hit compound appeared to elicit similar effects, rather than higher concentrations causing proportionally lower collagen synthesis levels. The exact collagen protein homolog was not distinguishable from the data; however, higher absorbance values corresponded to higher levels of gross collagen biosynthesis in the cell sample. All results were compared to the value of the positive control for each respective experiment to determine statistical differences between the groups. Comparisons to the negative control groups were not evaluated individually because this group was simply used to indicate the presence of collagen molecules and not other structural proteins.

To effectively evaluate the impact of the different hit compounds on Npas2 expression and total collagen deposition in the wound site, Figure 2 displays total collagen amounts for day 7 and day 14 while Figure 3 only shows results from day 7. This discrepancy was due to the observation in Figure 2, the collagen

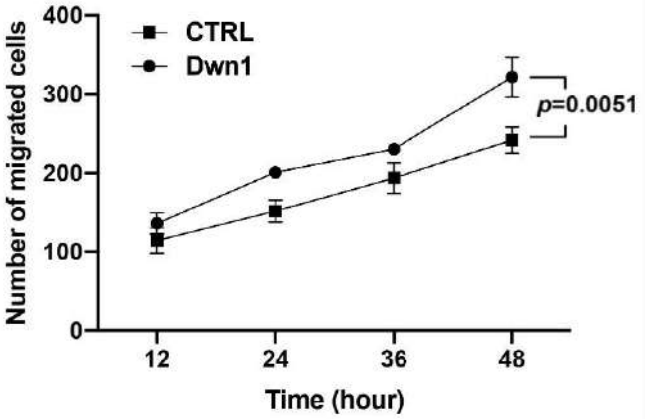


Figure 7. Dwn1 Treated Cell Migratory Assay Results Plotted. The number of migratory cells was recorded throughout 48-hour treatment for both control and Dwn1 treated groups. The average number of Dwn1 treated cells indicated significantly higher cell migration beyond the scratch test border for all time points. The p-value was calculated between these groups (p = 0.0051) . Graphs represent the average ± STD.

deposition values on day 7 were relatively equal, but there were significant decreases in collagen deposition for Dwn1 and Dwn5 on day 14. Consequently, the day 7 data showed that the induced effect by the hit compounds was a gradual decrease over the course of the wound healing timeline. By doing this, the hit compounds can maintain a healthy level of collagen in the wound site during the preliminary healing stages to facilitate rapid wound closure and cellular proliferation (3). Since the formation of HTS does not develop until late stages of the wound healing process (6), Figure 2 indicated that the therapeutic downregulation of Npas2 by the hit compounds may alter late-stage wound healing mechanisms more directly than early-stage, which would lead to shorter overall recovery times and an improvement in postoperative success (8). The data presented in Figure 3 is important because despite there being significant downregulation of C1 or C3 synthesis genes for Dwn1 and Dwn5 on day 7, there was no change in the total collagen biosynthesis. Figures 2 and 3 can be taken together to show that due the body's native tissue response to injury, downregulation of collagen synthesis genes causes a phenotypic change in the wound healing process during the later stages when HTS formation is likely to occur rather than in the preliminary stages of establishing a collagen scaffold for cell proliferation. This later activation of collagen suppression serves to reduce the over-expression of collagen, a phenomenon encountered in the development of many kinds of scars aside from HTS (13). These scars can lead to significant damage to a patient's psychological wellbeing, decreased satisfaction with life, altered perception of body image, and higher rates of post-traumatic stress disorder, alcoholism, imprisonment, unemployment, or marital discord (17). In summary, the data presented shows that although there appears to be no significant change in total collagen content on day 7, likely due to the native injury response, the effects of the hit compounds are still able to limit the formation of HTS during later stages of healing.

As shown in Figures 6 and 7, treatment by Dwn1 induced faster cell migration towards a wound site allowing Dwn1 to speed up the cell migration and decrease the overall wound healing time. The ability for cells to infiltrate the wound area more quickly after initial injury has been shown to not only quicken the recovery process (14), but also more closely regulate the amount of collagen deposited into the wound site, thereby preventing the excessive deposition of collagen proteins (11). Figure 5 shows that Dwn1 was able to significantly downregulate Npas2 activity throughout 48 hours of initial treatment. As such, Dwn1 seems to be a promising candidate for clinical applications due to its ability to regulate gene expression levels throughout multiple days in the wound healing process.

Although the study set out to determine the ability of the five different compounds to regulate Npas2 activity, only Dwn1 and Dwn5 emerged as being able to significantly affect overall collagen biosynthesis through regulation of collagen synthesis genes. These compounds offer the possibility to serve as therapeutic agents for dermal fibroblasts. The other three molecules (Dwn2, Dwn3, and Dwn4) did not exhibit sufficient downregulation of collagen, thereby preventing them from being used as possible treatments for HTS in a clinical setting for this and future studies. Although the exact cell signaling pathway taken by either of these two compounds is currently unknown, the results from these experiments set forth the possibility for these two compounds to follow a similar pathway within the cell. Future directions for this research include in vivo testing of mouse incisional models for treatment with Dwn1 and Dwn5. Future studies could be used to identify the specific pathways used by Dwn1 and Dwn5 to alter gene activity and provide insights into a possible signal cascade that can be utilized for more extensive wound healing studies. Consequently, supplementation with either of these two agents remains a promising possibility for incisional wound healing.

This work has the potential to broadly impact the field of wound healing, as it presents a novel method for the clinical control of the wound healing process. The ability to impact the speed of wound healing relies on specific circadian control. Keratinocyte re-epithelialization, an intermediate step in the wound healing process, has been shown to be under circadian control (5). Scratch tests conducted on keratinocyte monolayers showed higher cell mobility during the active phase of cellular metabolism (17). This coincides with the results established by Figures 6 and 7, in which higher rates of cell mobility indicate a faster rate of wound closure and, thus, wound healing. This work demonstrates the ability of Dwn1 and Dwn5 to attenuate the activity of fibroblasts. Myofibroblasts specifically have been shown to secrete excessive collagen into the ECM in response to injury (1) due to the presence of profibrotic cytokines capable of inducing the differentiation of myofibroblasts and production of excess collagen (8). As such, this study offers insights into how the regulation of these cell types can be influenced by hit compound treatment, thus supporting the capability of these compounds to be useful in a clinical setting.

The field of wound healing is progressing toward the use of novel agents for clinical therapeutics in order to help facilitate the wound healing process while mitigating both patient and fabrication costs (18). These compounds can better modulate cellular behavior in response to injury. This work has direct implications in the healing of surgical wounds, specifically following plastic surgery, as the ability to prevent the formation of HTS will enhance both clinical success and patient experience (17). Patients with HTS following craniofacial surgery not only exhibit poorer clinical success following the procedure, but also often undergo subsequent procedures to remove or hide the scar itself (19). In doing so, they risk alterations to the original work, worsening of scar presentation, and/or damage to surrounding healthy tissue (20). Each year, over 100 million people acquire some type of scar, with 55 million of those individuals paying for scar revision procedures that cost as much as $4,000 (2). In total, this accounts for $220 million in scar revision procedures alone, with many of these procedures being paid by the individual, rather than insurance companies. This work has broader implications in the ability for millions of individuals worldwide to save both time and money in their healthcare costs by accelerating the wound healing process and avoiding the development of HTS.

REFERENCES

1. H. Sasaki et al., Neuronal pas domain 2 (npas2)deficient fibroblasts accelerate skin wound healing and dermal collagen reconstruction. Anat Rec. 303, 1630–1641 (2019), doi:10.1002/ar.24109.
2. D. Wolfram, A. Tzankov, P. Pülzl, H. Piza-Katzer, Hypertrophic scars and keloids—a review of their pathophysiology, risk factors, and therapeutic management. Deramtol Surg. 35, 171–181 (2009), doi:10.1111/j.1524-4725.2008.34406.x.
3. D. W. Friedman et al., Regulation of collagen gene expression in keloids and hypertrophic scars. J Surg Res. 55, 214–222 (1993), doi:10.1006/jsre.1993.1132.
4. J. A. Garcia et al., Impaired cued and contextual memory in NPAS2-deficient mice. Science. 288, 2226–2230 (2000), doi:10.1126/science.288.5474.2226.
5. N. Boyera, I. Galey, B. A. Bernard, Effect of vitamin C and its derivatives on collagen synthesis and cross-linking by normal human fibroblasts. Int J Cosmetic Sci. 20, 151–158 (1998), doi:10.1046/j.1467-2494.1998.171747.x.
6. M. K. Gordon, R. A. Hahn, Collagens. Cell Tissue Res. 339, 247–257 (2009), doi:10.1007/s00441-009-0844-4.
7. H. Kuivaniemi, G. Tromp, Type III collagen (COL3A1): Gene and protein structure, tissue distribution, and associated diseases. Gene. 707, 151–171 (2019), doi:10.1016/j.gene.2019.05.003.
8. N. Meisler et al., Glucocorticoids coordinately regulate type I collagen PROA1 promoter activity through both the glucocorticoid and transforming growth factor β response elements: A novel mechanism of glucocorticoid regulation of eukaryotic genes. J Cell Biochem. 59, 376–388 (1995), doi:10.1002/jcb.240590309.
9. R. A. Akhtar et al., Circadian cycling of the mouse liver transcriptome, as revealed by cdna microarray, is driven by the suprachiasmatic nucleus. Curr Bio. 12, 540–550 (2002), doi:10.1016/s0960-9822(02)00759-5.
10. D. O'Neil et al., Dysregulation of NPAS2 leads to altered metabolic pathways in a murine knockout model. Mol Genet Metab. 110, 378–387 (2013), doi:10.1016/j.ymgme.2013.08.015.
11. N. P. Hoyle et al., Circadian actin dynamics drive rhythmic fibroblast mobilization during wound healing. Sci Transl Med. 9 (2017), doi:10.1126/scitranslmed.aal2774.
12. C. Bertolucci et al., Evidence for an overlapping role of clock and NPAS2 transcription factors in liver circadian oscillators. Mol Cell Bio. 28, 3070–3075 (2008), doi:10.1128/mcb.01931-07.
13. S. A. Brown, Circadian clock-mediated control of stem cell division and differentiation: Beyond night and day. Development. 141, 3105–3111 (2014), doi:10.1242/dev.104851.
14. E. J. Cable, K. G. Onishi, B. J. Prendergast, Circadian rhythms accelerate wound healing in female siberian hamsters. Physiology & Behavior. 171, 165–174 (2017), doi:10.1016/j.physbeh.2016.12.019.
15. J. Rutter, M. Reick, L. C. Wu, S. L. McKnight, Regulation of clock and NPAS2 DNA binding by the redox state of Nad Cofactors. Science. 293, 510–514 (2001), doi:10.1126/science.1060698.
16. A. J. Bailey et al., Characterization of the collagen of human hypertrophic and normal scars. Biochim Biophys Acta - Protein Struct. 405, 412–421 (1975), doi:10.1016/0005-2795(75)90106-3.
17. G. Krishna, G. Gopalakrishnan, S. Goel, In vitro and in vivo genotoxicity assessment of the dopamine receptor antagonist molindone hydrochloride. Environ Mol Mutagen. 57, 288–298 (2016), doi:10.1002/em.22007.
18. S. N. Haque et al., "Effects of BMAL1 Manipulation on the Brain's Master Circadian Clock and Behavior." Yale J Bio Med. 92 251-258 (2019).
19. Y. Shibuya et al., Therapeutic downregulation of neuronal pas domain 2 (NPAS2) promotes surgical skin wound healing. eLife. 11 (2022), doi:10.7554/elife.71074.
20. J. Duan, E. N. Greenberg, S. S. Karri, B. Andersen, The circadian clock and diseases of the skin. FEBS Lett. 595, 2413–2436 (2021), doi:10.1002/1873-3468.14192.

Design of a Brønsted acid catalyst for asymmetric Ritter reactions

Chris La[1], Vladislav A. Roytman, PhD[2], Fransisco Dean Toste, PhD[2]

[1]Department of Chemistry and Biochemistry. [2]Department of Chemistry, University of California, Berkeley.

ABSTRACT

The Ritter reaction describes an acid-catalyzed nucleophilic substitution of an alkene or alcohol with nitrile to produce an amide. Advances in the Ritter reaction have enabled its application in the synthesis of various molecules, including polymers and bioactive compounds. However, such applications are limited by the reaction's harsh acidic conditions and lack of asymmetric, or enantioselective, methods. While Ritter reactions have been catalyzed in milder conditions, no enantioselective catalytic methods exist in the literature. In this study, chiral 1,1'-bi-2-naphthol-derived Brønsted acids, specifically phosphoric, dithiophosphoric, and N-triflyl phosphoramide acids, were computationally investigated using density-functional theory studies (B3LYP/6-31+G*) as potential catalysts for the Ritter reaction since these acids provide both the acidity needed for the reaction to proceed and a scaffold for enantioinduction. Two benzylic alcohols, one with an additional *ortho*-hydroxy group, were used as the substrates in the example system, reacting with either acetonitrile or benzonitrile as the solvent electrophile. The resulting data indicated that in the stereodetermining step of these reactions, the transition states exhibited moderate enantioselectivity for the amide product with 0.1 to 3.9 kcal/mol difference in transition state energies, predicting up to 99% enantiomeric excess for a given Ritter reaction of substrate with a nitrile. The enantioselectivity of the Ritter reaction was hypothesized to be controlled by a suite of noncovalent interactions between the substrate complex and the catalyst. Interactions between the nitrile and the aryl substituents at the 3,3'-positions of the 1,1'-bi-2-naphthol backbone were identified to increase enantioselectivity via noncovalent interaction analysis. Future work for this project includes investigating more N-triflyl phosphoramides with modified aryl groups at the 3,3'-positions.

INTRODUCTION

The Ritter reaction is the chemical reaction of an alkene or an alcohol with a nitrile, which is any molecule with a carbon-nitrogen triple bond group. This reaction produces an amide in the presence of an acid; the general product structure is shown in Figure 1 (*1*). An example of the mechanism for the Ritter reaction for 1-phenylethanol and acetonitrile can be seen in Figure 2 (*2*). The key process of the Ritter reaction is an SN1 reaction of the nitrile on the alkene or alcohol substrate. This substitution occurs through a stepwise mechanism where a carbocation intermediate forms, allowing for the stereochemistry of the starting material to undergo inversion, where stereochemistry changes, or retention, where stereochemistry does not change (*3*).

There are many applications of the Ritter reaction, as amides are versatile intermediates in organic synthesis and compounds of interest in natural product synthesis. The most common industrial application of the Ritter reaction is the production of *tert*-octylamine, which is used as a component in resins and light stabilizers (*4*). The Ritter reaction has also been used in the synthesis of bioactive molecules, polymers, and pharmaceutical intermediates (*1*).

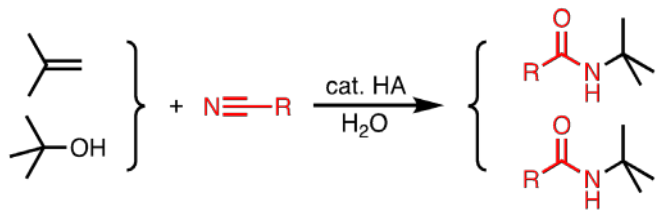


Figure 1. The Ritter reaction. An alkene or alcohol reacts with nitrile in the presence of an acid to produce an amide. An amide group is added to the non-terminal carbon of an alkene or the alpha carbon of an alcohol.

Although the Ritter reaction has been proven useful in these applications, the reaction itself contains parameters that limit its efficiency as compared to other amide synthesis reactions. The Ritter reaction generally requires stoichiometric, or equivalent, amounts of strong acid to proceed. This results in strongly acidic experimental conditions that limit starting materials for the Ritter reaction to only those that do not decompose in acid (*2*). Another issue of the Ritter reaction lies in its lack of enantioselectivity: an asymmetric version of the Ritter reaction has yet to be reported (*1*). Enantioselectivity refers to the formation of a specific enantiomer, or non-superimposable

mirror image of a given molecule, over another (*5*). Enantioselectivity plays an especially important role in the pharmaceutical industry since different enantiomers can have drastically different effects on the body (*6*). Although the stereochemistry of the substitution can be controlled, the instability of the carbocation intermediate (Figure 2, state [2]) makes it difficult to control stereoselectivity. However, the presence of this carbocation intermediate provides an approach to make the reaction stereoselective because the direction of nucleophilic attack of the nitrile (Figure 1, state [3]) can be controlled through enantioinduction. One strategy for enantioselective synthesis is the use of chiral catalysts, where one enantiomeric transition state will have more favorable, or stabilizing, interactions with the chiral ligands of the catalyst in comparison to the other (*7*). Approaches to designing chiral catalysts generally involve ligands that are large in size and/or contain multiple functional groups, as these increase noncovalent interactions between the catalyst and substrate (*7*).

To overcome the issue of acid stoichiometry, certain strong Brønsted acids can act as acid catalysts for the Ritter reaction. Acids such as *o*-benzenedisulfonimide and 2,4-dinitrobenzenesulfonic acid have successfully catalyzed the Ritter reaction (*8*, *9*). When present in mild, sub-stoichiometric amounts, these acid catalysts still allow the Ritter reaction to proceed with high product yield. Based on previous results with sulfonic-type acids, it is possible that chiral 1,1'-bi-2-naphthol (BINOL)-derived acids would also work for the Ritter reaction, as these species are commonly used as Brønsted acid catalysts in asymmetric reactions (*7*). The axial chirality of the BINOL backbone has the potential to induce enantioselectivity, which could lead to the development of an asymmetric Ritter reaction. In this work, chiral BINOL-derived Brønsted acids will be investigated through density-functional theory energy calculations as possible catalysts for asymmetric Ritter reactions.

MATERIALS AND METHODS

Identifying acids as prospective catalysts

All catalyst candidates contained a BINOL scaffold with phenyl groups located at the 3,3'-positions of the BINOL backbone. The potential catalysts investigated in this study were three acids derived from (*S*)-BINOL: phosphoric acid, dithiophosphoric acid, and N-triflyl phosphoramide (NTP) acid (Figure 3). NTP acids have been previously studied as chiral acid catalysts in other synthesis reactions (*7*). Moreover, the NTP acid was expected to be the most acidic of all these designs and has a similar pKa value as *o*-benzenedisulfonimide, which is an established Ritter catalyst (*7*).

The performance of each of these acids as catalysts in the Ritter reaction was theoretically investigated through density-functional theory (DFT) calculations in Gaussian 16, an electronic structure modeling program. DFT is a computational quantum mechanical modeling method that can be used to predict the behavior of atomic systems. In this investigation, DFT was used to determine the energetics of reaction states (*10*). The

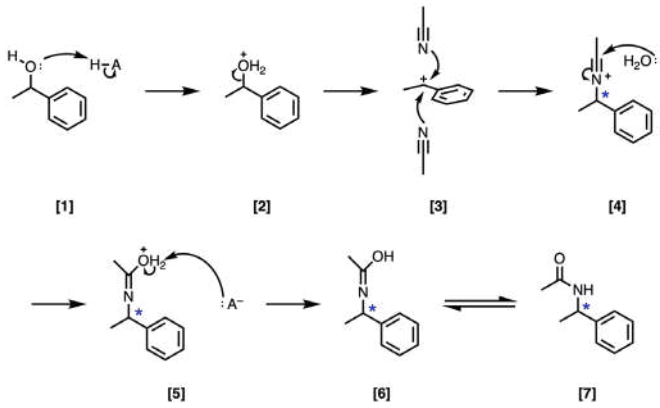


Figure 2. Mechanism of the Ritter reaction, shown for 1-phenylethanol and acetonitrile. The acid and its conjugate base is represented by HA and A–, respectively. In state [3], the nitrile can perform a nucleophilic attack on the carbocation from either direction. The stereocenters of the reaction are marked with a star. The presence of a star after state [3] indicates that it is the stereodetermining step of the reaction mechanism. Two enantiomers of the final amide product can form.

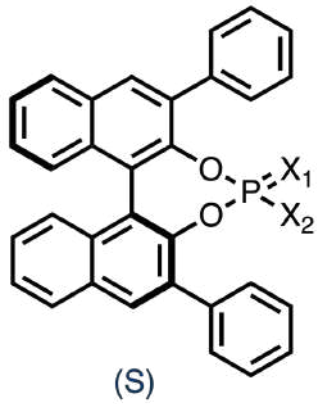
specific parameters used in Gaussian 16 included B3LYP as the functional, 6-31+G* as the basis, and GD3 as the empirical dispersion model. As shown in state [1] of Figure 2, a proton transfer from the acid to the substrate occurs as the initial step of the Ritter reaction. The ability of each acid to facilitate this proton transfer was used to determine its suitability as a catalyst in the Ritter reaction.

Enantioinduction studies

Acids determined to be sufficiently catalytic were then studied for enantioinduction using the same computational parameters. Acetonitrile and benzonitrile were modeled as the reaction solvent using an additional parameter: the polarizable continuum model, which is commonly used to model solvation effects. Choosing acetonitrile or benzonitrile as the reaction solvent reflected real experimental conditions of the Ritter reaction, where the source of nitrile usually is the solvent.

The substrates investigated were 1-phenylethanol and 1-phenylethanol with an *ortho*-hydroxy group, whereas the reacting nitriles were acetonitrile and benzonitrile. The extra alcohol group in the *ortho*-hydroxy species was suspected to lower activation energies via resonance stabilization of the carbocation (*11*).

For each set of reactants, two enantiomeric, stereodetermining transition states were determined: one that led to the (*S*)-amide product and one that led to the (*R*)-amide product. These transition states occur between states [3] and [4] of Figure 2. DFT calculations were performed to determine the energetics and intermolecular interactions of these transition states. The transition state activation energies were calculated relative to the energy of the isolated starting materials for the reaction. The energy calculations also indicated the energetic favorability, or feasibility, of any of the given catalysts to proceed in the Ritter reaction. The lower-energy transition state will yield the favored enantiomer because lower-energy pathways are generally favored in



Catalyst	X_1	X_2
Phosphoric	O	OH
Dithiophosphoric	S	SH
N-triflyl phosphoramide	O	NHTf

Figure 3. The three acids proposed as catalysts for the Ritter reaction. All catalyst candidates contained a 1,1′-bi-2-naphthol (BINOL) scaffold with phenyl groups located at the 3,3′-positions of the BINOL backbone. A line diagram of the general structure of the acids are shown on the left. The table on the right shows the name of each acid, with the atoms or groups at the X1 and X2 positions of the acid also shown. All catalysts have (*S*) axial chirality.

reactions. Likewise, a greater difference in energy between the two enantiomeric transition states leads to a greater degree of enantioselectivity. Noncovalent interaction analysis was used to determine the key noncovalent interactions influencing enantioselectivity.

In addition to the NTP acid shown in Figure 3, another NTP acid was also investigated, where 2,4,6-triisopropylphenyl groups replaced the phenyl groups at the 3,3'-positions on the BINOL backbone. It was expected that more noncovalent interactions will occur between the catalyst and the substrate due to the isopropyl groups located along the phenyl ring of this aryl group, which would lead to more enantioselectivity. Due to the increased complexity and size of the atomic system with these modified aryl groups, only the case of the 1-phenylethanol with an *o*-hydroxy group reacting with acetonitrile was investigated.

RESULTS

NTP acids as prospective catalysts facilitate initial proton transfer of the Ritter reaction

In the calculations involving protonated 1-phenylethanol and a deprotonated acid, optimizations showed that a proton transfer occurred between the two species for the phosphoric and dithiophosphoric acid. In other words, the charged species of state [2] of Figure 1 for the phosphoric and dithiophosphoric acid reverted to the neutral species of state [1] of Figure 1. For the same calculation with the NTP acid, this reverse proton transfer did not occur. In the cases of phosphoric and dithiophosphoric acid, there was no energy barrier for the charged species to undergo reverse proton transfer back to neutral species, which indicated that less substrate is present in the charged state at equilibrium and there would be less protonated substrate to proceed with the reaction. In the case of the NTP acid, the presence of an energy barrier for reverse proton transfer meant that more substrate was present in the charged state at equilibrium. Based on these results, efforts were then focused on investigating the transition state energetics of the Ritter reaction in the presence of the NTP acid.

Ritter reactions in the presence of an NTP acid have enantioselective energy differences between stereodetermining transition states

Four pairs of enantiomeric, stereodetermining transition states in the presence of the NTP acid were calculated: 1-phenylethanol reacting with acetonitrile, 1-phenylethanol reacting with benzonitrile, 1-phenylethanol with an *ortho*-hydroxy group reacting with acetonitrile, and 1-phenylethanol with an *ortho*-hydroxy group reacting with benzonitrile.

The transition state energies of 1-phenylethanol reacting with acetonitrile are shown in Figure 4A. The difference in transition state energies was $\Delta(\Delta G‡)$ = 0.1 kcal/mol, which led to a product ratio of (R:S) = 1.17:1 in favor of the (*S*)-amide product. For 1-phenylethanol reacting with benzonitrile, the difference in transition state energies was $\Delta(\Delta G‡)$ = 1.1 kcal/mol, which led to a product ratio of (R:S) = 5.93:1 in favor of the (*R*)-amide product (Figure 4B). In the case of 1-phenylethanol with an *o*-hydroxy group reacting with acetonitrile, the difference in transition state energies was $\Delta(\Delta G‡)$ = 0.7 kcal/mol, which generated a product ratio of (R:S) = 3.04:1 in favor of the (*R*)-amide product (Figure 4C). The transition state energies of 1-phenylethanol with an *o*-hydroxy group reacting with benzonitrile are shown in Figure 4D. The difference in transition state energies was $\Delta(\Delta G‡)$ = 0.6 kcal/mol, which resulted in a product ratio of (R:S) = 2.60:1 in favor of the (*R*)-amide product.

Modified aryl groups at the 3,3′-positions of the BINOL backbone led to a greater energy difference between stereodetermining transition states

The transition state energies of 1-phenylethanol with an *o*-hydroxy group reacting with acetonitrile in the presence of the modified NTP acid are shown in Figure 4E. The difference in transition state energies is $\Delta(\Delta G‡)$ = 3.9 kcal/mol, which led to a product ratio of (R:S) = 783:1 in favor of the (*R*)-amide product.

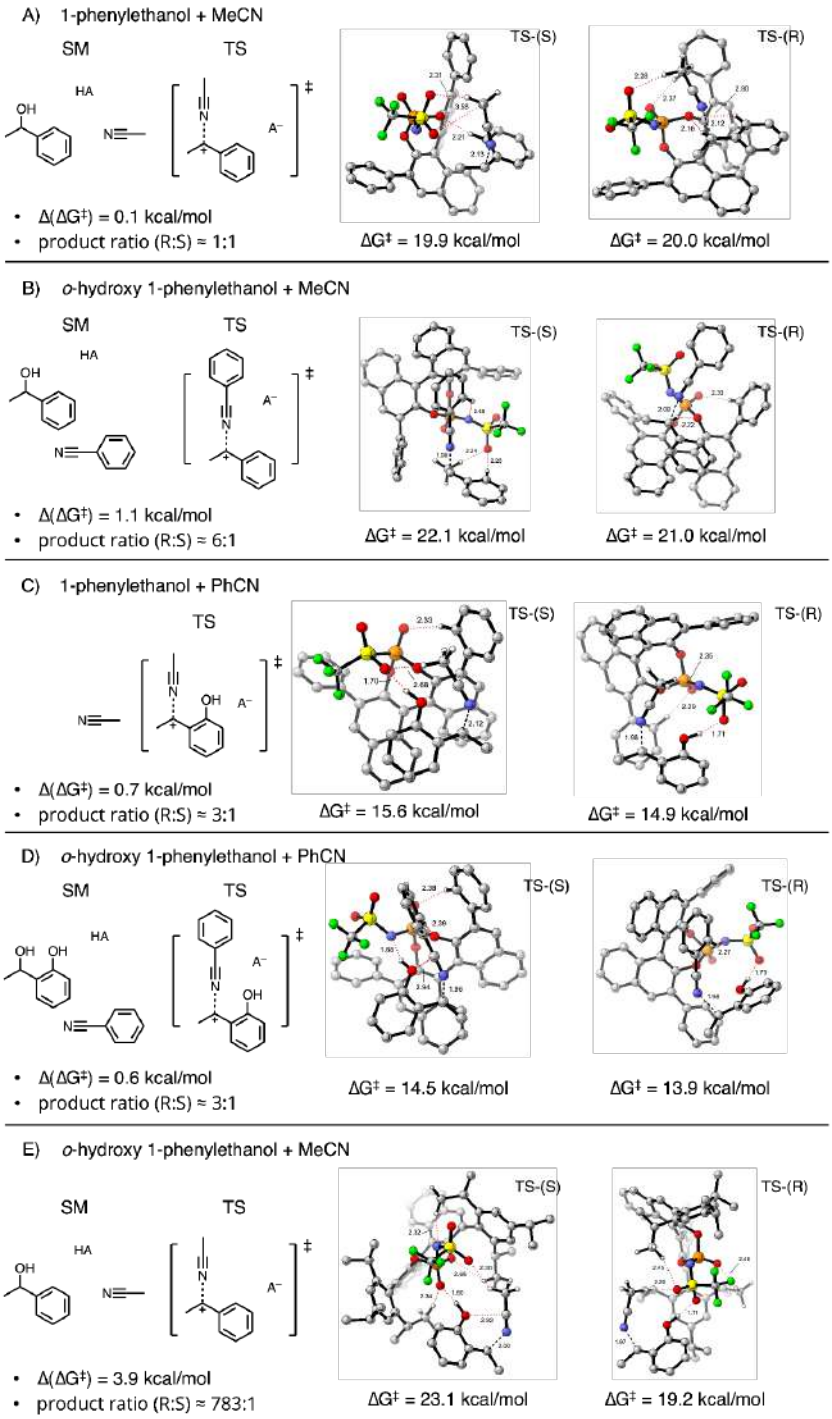


Figure 4. Enantiomeric transition states of the various substrates and nitriles in the presence of an N-triflyl phosphoramide acid. Line diagrams of the species are shown on the left, and the molecular structures of each transition state are shown on the right. Key noncovalent interactions are represented by red-dotted bonds with distances in Å. The difference in transition state energies is represented by Δ(ΔG‡). **A)** Transition states of 1-phenylethanol reacting with acetonitrile. The difference in transition state energies were Δ(ΔG‡) = 0.1 kcal/mol, which led to a product ratio of (R:S) = 1.17:1 in favor of the (S)-amide product. **B)** Transition states of 1-phenylethanol reacting with benzonitrile. The difference in transition state energies were Δ(ΔG‡) = 1.1 kcal/mol, which led to a product ratio of (R:S) = 5.93:1 in favor of the (R)-amide product. **C)** Transition states of 1-phenylethanol with an o-hydroxy group reacting with acetonitrile. The difference in transition state energies was Δ(ΔG‡) = 0.7 kcal/mol, which led to a product ratio of (R:S) = 3.04:1 in favor of the (R)-amide product. **D)** Transition states of 1-phenylethanol with an o-hydroxy group reacting with benzonitrile. The difference in transition state energies was Δ(ΔG‡) = 0.6 kcal/mol, which led to a product ratio of (R:S) = 2.60:1 in favor of the (R)-amide product. E) Transition states of 1-phenylethanol with an o-hydroxy group reacting with acetonitrile, in the presence of the modified N-triflyl phosphoramide acid. The difference in transition state energies was Δ(ΔG‡) = 3.9 kcal/mol, which led to a product ratio of (R:S) = 783:1 in favor of the (R)-amide product. Abbreviations: HA/A–, N-triflyl phosphoramide acid and its conjugate base (A-D have phenyl groups at the 3,3′-positions on the BINOL backbone, E has 2,4,6-triisopropylphenyl groups at the 3,3′-positions on the BINOL backbone); SM, starting materials; TS, transition state.

DISCUSSION

Covalent bonds should not form or break unless there is no energy barrier for bond rearrangement; thus, the proton transfer that occurred for the deprotonated phosphoric and dithiophosphoric acids upon optimization implied that there was no energy barrier for either direction of proton transfer. The absence of an energy barrier allowed for decomposition back to the starting materials once the initial protonation occurs. However, since no proton transfer occurred in the presence of the deprotonated NTP acid, there would likely be less decomposition back to the starting materials after the initial protonation. Thus, NTP acids were determined to be the most suitable functional group for use as a catalyst in the Ritter reaction.

Of the transition states that were calculated in the presence of NTP acid (Figures 4A-D), the (*R*)-amide product was favored in all except in the case of 1-phenylethanol reacting with acetonitrile, where the (*S*)-amide product was favored and the difference in transition state energies was the lowest. The inconsistency in preference of enantiomeric product likely indicates that the calculation for 1-phenylethanol reacting with acetonitrile must be revised since all the amide products are expected to have the same favored stereochemistry in the presence of the same chiral catalyst.

The key noncovalent interactions stabilizing the transition states and thereby influencing the enantioselectivity of the reaction consisted of a suite of C–H···O and CH/π interactions between the catalyst and the substrate complex. In the case of the alcohol substrates with *o*-hydroxy groups, hydrogen bonding was also involved. It was suspected that benzonitrile would lead to greater enantioselectivity due to the potential of more CH/π and possibly π /π interactions. However, based on the inconsistent results when comparing transition states with acetonitrile and benzonitrile, no conclusions could be drawn about the effect of benzonitrile on enantioselectivity in the Ritter reaction. In the case where 1-phenylethanol was modeled as the substrate, using benzonitrile as the nucleophile instead of acetonitrile led to an increase in the difference in transition state energies and an increase in enantioselectivity. In the case where 1-phenylethanol with an *o*-hydroxy group was modeled as the substrate, using benzonitrile as the nucleophile instead of acetonitrile only resulted to a slight decrease in the difference in transition state energies and a slight decrease in enantioselectivity.

Comparison of the transition states of the two alcohol substrates revealed that the addition of *o*-hydroxy groups did not have a clear effect on transition state energy difference. When acetonitrile was modeled as the reaction solvent, having 1-phenylethanol with an *o*-hydroxy group as the substrate (Figure 4C) instead of 1-phenylethanol (Figure 4A) caused an increase in the difference in transition state energies and thus an increase in enantioselectivity. In contrast, when benzonitrile was modeled as the reaction solvent, having 1-phenylethanol with an *o*-hydroxy group as the substrate (Figure 4D) instead of 1-phenylethanol (Figure 4B) resulted in a decrease in the difference in transition state energies and a decrease in enantioselectivity. However, the absolute difference in energy relative to the isolated starting materials was lower for all cases involving substrates with the *o*-hydroxy groups. By lowering the activation energy barrier, the Ritter reaction can be performed at a practical rate even at a lower reaction temperature. Lower reaction temperatures preserve the selectivity that arises from transition state energy differences, thus optimizing experimental conditions for enantioselectivity.

The modification of the NTP acid led to an increased transition energy difference and thus, increased enantioselectivity (Figure 4E). The 2,4,6-triisopropylphenyl groups were seemingly more involved in the noncovalent interactions that stabilized the enantiomeric transition states compared to the phenyl groups.

For all the transition states calculated, the energy difference between the enantiomeric, stereodetermining transition states ranges from 0.1 to 3.9 kcal/mol, predicting up to 99% enantiomeric excess (ee). Enantiomeric excess refers to the excess of one enantiomer over the other in a mixture of enantiomers and can be expressed as the percent difference between the major enantiomer and the minor enantiomer (*12*). A larger difference in transition state energies leads to a greater %ee, which indicates greater enantioselectivity.

CONCLUSION

The synthesis of amides is a contemporary challenge in organic synthesis. Though the functional group is prevalent in pharmaceuticals and bioactive compounds, current methods for amide formation generally require stoichiometric quantities of activating reagents, which tend to be expensive and wasteful (*13*). Developing catalytic approaches for amide synthesis, such as an asymmetric Ritter reaction, would improve the atom economy of the process. Such approaches could likewise help mitigate the impact of the pharmaceutical and other industries on the environment.

Although the Ritter reaction is useful in the synthesis of amides and has the potential to amend the sustainability issues involved in amide synthesis, its limitations lie in its acidity requirements and lack of enantioselectivity. An NTP acid was investigated through DFT modeling and has been determined to be a theoretically suitable catalyst for asymmetric Ritter reactions. The chiral scaffold of the acid enables enantioinduction and provides stabilizing interactions for the transition state. Key noncovalent interactions that stabilized each transition state were identified, and no consistent effect on enantioselectivity could be determined upon modifying the reacting nitrile. However, adding an *o*-hydroxy group onto the 1-phenylethanol substrate lowered the absolute difference in energy relative to the isolated starting materials, which optimizes experimental conditions for enantioselectivity. Modifying the aryl groups at the 3,3'-positions of the BINOL backbone also led to increased stabilizing interactions, which resulted in increased enantioselectivity.

Future work for this project includes optimizing the enantioselectivity of the catalyst through further modifications of the

aryl groups at the 3,3'-positions of the BINOL backbone. 2,4,6-triisopropylphenyl groups were already investigated, but other groups such as triphenylsilyl groups also have the potential to induce more enantioselectivity. In addition to the noncovalent interaction analyses that were performed for each transition state, distortion interaction analyses are also being explored.

REFERENCES

1. G. M. Ziarani, F. S. Hasankiadeh, F. Mohajer, Recent Applications of Ritter Reactions in Organic Syntheses. *ChemistrySelect*. **5**, 14349–14379 (2020). doi: 10.1002/slct.202003470
2. A. Guérinot, S. Reymond, J. Cossy, Ritter Reaction: Recent Catalytic Developments. *Eur. J. Org. Chem.* **2012**, 19–28 (2012). doi: 10.1002/ejoc.201101018
3. C. Kelk Ingold, E. Rothstein, CLVI.—The nature of the alternating effect in carbon chains. Part XXV. The mechanism of aromatic side-chain substitution. *J. Chem. Soc. (Resumed).* **0**, 1217–1221 (1928). doi: 10.1039/JR9280001217
4. D. Jiang, T. He, L. Ma, Z. Wang, Recent developments in Ritter reaction. *RSC Adv.* **4**, 64936–64946 (2014). doi: 10.1039/C4RA10784E
5. G. P. Moss, Basic terminology of stereochemistry (IUPAC Recommendations 1996). *Pure Appl. Chem.* **68**, 2193–2222 (1996). doi: 10.1351/pac199668122193
6. J. McConathy, M. J. Owens, Stereochemistry in Drug Action. *Prim. Care Companion J. Clin. Psychiatry.* **5**, 70–73 (2003). doi: 10.4088/pcc.v05n0202
7. D. Parmar, E. Sugiono, S. Raja, M. Rueping, Complete Field Guide to Asymmetric BINOL-Phosphate Derived Brønsted Acid and Metal Catalysis: History and Classification by Mode of Activation; Brønsted Acidity, Hydrogen Bonding, Ion Pairing, and Metal Phosphates. *Chem. Rev.* **114**, 9047–9153 (2014). doi: 10.1021/cr5001496
8. M. Barbero, S. Cadamuro, S. Dughera, G. Ghigo, o-Benzenedisulfonimide and its chiral derivative as Brønsted acids catalysts for one-pot three-component Strecker reaction. Synthetic and mechanistic aspects. *Org. Biomol. Chem.* **10**, 4058–4068 (2012). doi: 10.1002/ejoc.200800931
9. S. Lin, E. N. Jacobsen, Thiourea-catalysed ring opening of episulfonium ions with indole derivatives by means of stabilizing non-covalent interactions. *Nature Chem.* **4**, 817–824 (2012). doi: 10.1038/nchem.1450
10. R. O. Jones, Density functional theory: Its origins, rise to prominence, and future. *Rev. Mod. Phys.* **87**, 897–923 (2015). doi: 10.1103/RevModPhys.87.897
11. F. Göricke et al., Phosphoric Acid Catalyzed Formation of Hydrogen-Bonded *o* -Quinone Methides. Enantioselective Cycloaddition with β-Dicarbonyl Compounds toward Benzannulated Oxygen Heterocycles. *J. Org. Chem.* **85**, 11699–11720 (2020). doi: 10.1021/acs.joc.0c01375
12. R. E. Gawley, Do the Terms "%ee" and "%de" Make Sense as Expressions of Stereoisomer Composition or Stereoselectivity? *J. Org. Chem.* **71**, 2411–2416 (2006). doi: 10.1021/jo052554w
13. M. T. Sabatini, L. T. Boulton, H. F. Sneddon, T. D. Sheppard, A green chemistry perspective on catalytic amide bond formation. *Nat. Catal.* **2**, 10–17 (2019). doi: 10.1038/s41929-018-0211-5

Oxygen reduction species in ambient air methane to methanol conversion

David M. Dumas[1], Benjamin S. Natinsky[1], Chong Liu[1]

[1]Department of Chemistry and Biochemistry.

ABSTRACT

Methane is an energy-rich byproduct of gas flaring, oil extraction, and various other industrial practices. Although methane is energy-rich, its gaseous state of matter complicates transportation, so converting methane to a liquid fuel such as methanol would provide access to a large reservoir of energy-rich chemicals for organic synthesis. Such a conversion pathway exists, but researchers lack the necessary insight on specific mechanistic steps such as the identity of the primary terminal oxidant. In the conversion, a negative voltage is applied to an electrolyte solution, simultaneously activating metalloradical catalysts (rhodium porphyrin) and generating an oxygen-free zone. The silicon nanowire electrodes activate rhodium via electrochemical reduction to form a methylated-complex with methane, which ultimately reacts with a reduced oxygen species to produce methanol. These experiments study oxygen reduction reactions, which generate potentially green oxidants prevalent in industry, research, and nature. Electron paramagnetic resonance and ultraviolet-visible absorption spectroscopy tests verified that superoxide was electrochemically generated during the conversion of methane to methanol. However, non-electrochemical experiments between the proposed reaction intermediates suggested that hydroperoxide was the primary terminal oxidant as opposed to superoxide. Uncovering the mechanism of this conversion would help understand oxygen reduction reactions, optimize existing processes, and formulate greener pathways involving photoelectrochemistry.

INTRODUCTION

Converting methane (CH_4) to a commodity chemical like methanol (CH_3OH) using oxygen (O_2) would finally allow industry to access the immense reservoirs of CH_4 built up over decades of waste disposal, oil extraction, and decomposition (*1–3*). Current chemical conversion technologies rely on high temperatures and pressures to facilitate the heterogeneous conversion of gaseous CH_4 to liquid CH_3OH. This process, although efficient, is highly energy-intensive and requires a large amount of infrastructure to be implemented in current industrial facilities (*2*). Conversely, biological methods prove more favorable since they can be carried out in ambient air under less energy intensive conditions. However, to preserve the biocatalysts in this conversion, stringent requirements must be maintained, which can also have a high infrastructural demand (*2*). Establishing an electrochemical method to carry out this conversion would be ideal, as it would require less infrastructure to implement in existing industrial facilities, increasing its economic appeal (*1–4*). While the Liu group has already developed an electrochemical conversion, information regarding the terminal oxidation of CH_4 and the identity of the reactive oxygen species (ROS) requires further investigation to both optimize and properly scale this conversion process (*3,4*).

O_2 plays an important role as a reactant in a plethora of biological processes and offers limitless possibilities for scientists to functionalize different organic molecules (*5*). O_2-reduction reactions (ORRs) are regularly utilized in industry and research because the ROSs generated from ORRs can serve as potentially green oxidants in electrochemical reactions (*5*). Green oxidants represent a sustainable and largely non-toxic class of chemicals that may serve as safer reagents in previously hazardous syntheses (*5*). Superoxide (O_2^-) is one ROS candidate worthy of investigation largely because there are few reactions known to utilize $O_2^{\cdot-}$ and its general instability and short life span make it difficult to analyze (*5*). Investigating this ROS in the methane-to-methanol conversion can offer tremendous insight into $O_2^{\cdot-}$ reactions and potentially utilizing O_2^- as a future green oxidant (*5*). It is important to note that in the presence of free protons (H^+), $O_2^{\cdot-}$ decomposes to form hydroperoxide ($^{\cdot}OH$) and O_2 (*5*). This decomposition pathway indicates that even if $O_2^{\cdot-}$ is electrochemically generated, it may not be the terminal oxidant in this conversion.

Electron deficient organometallic catalysts are known to be strong C-H activators, but coupling the reaction with an oxidant can quench the catalytic species before reacting with CH_4 (*3,6*). Previous work has demonstrated that silicon nanowires (SiNWs), electricity, rhodium porphyrin catalysts (Rh-cat), and air can be used to generate CH_3OH (*3*). However, the complexity of this reaction lies in electrochemically sequestering

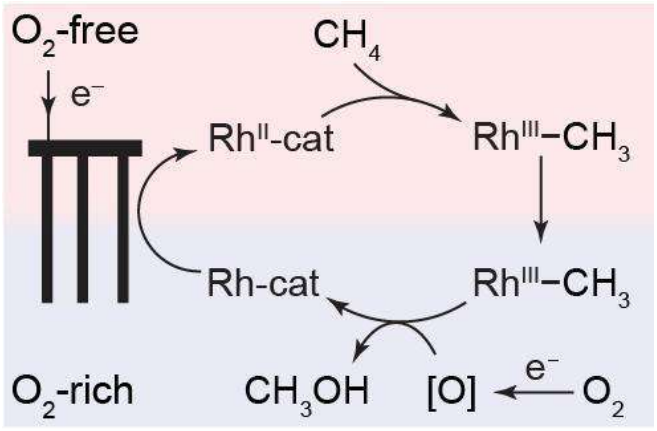


Figure 1. Proposed mechanism for methane-methanol conversion. The reaction of rhodium catalyst (Rh-cat), methane (CH_4) and a reactive oxygen species ([O]) is facilitated electrochemically to generate methanol. Silicon nanowires (SiNWs) serve as an electrode that activates the rhodium catalyst (RhII-cat) to react with methane and generates an oxygen (O_2)-gradient. The methylated rhodium complex (RhIII-CH_3) is then oxidized by the reactive oxygen species ([O]) generated by the SiNWs. This figure was taken from figure 1 of reference 3 with permission from the authors (3).

the catalyst from kinetically favorable side reactions (*3*). Application of cathodic potential reduces the Rh-cat to its C-H active form and simultaneously reduces O_2, which generates a local anaerobic region by depleting O_2 near the SiNW electrode (Figure 1) (*3,6*). This separation is required because Rh-cat is more reactive with O_2 than with CH4, so electrochemically generating an O_2-gradient within the SiNW apparatus allows the methylated rhodium complex (CH_3-Rh-cat) to form before free O_2 can interfere (*3,6*). The reactive oxygen species generated by the negative voltage is believed to be either superoxide or hydrogen peroxide (·OH) because they are known oxidants in ORRs (*3,5*). This study conducted several tests to identify the likely ROS generated electrochemically and used as a terminal oxidant in CH_4 functionalization.

Detecting the short-lived superoxide species can prove difficult without spin trapping agents and selective superoxide chromogens to isolate these unstable molecules for spectroscopic analysis (*4*). Electron paramagnetic resonance (EPR) spectroscopy can identify free radicals in the solution and distinguish between the presence of $O_2^{\cdot-}$ and .OH, thereby uncovering the primary ROS in electrolysis experiments (*5*). Nitro blue tetrazolium (NBT), a selective superoxide chromogen, serves as a color indicator that turns purple when superoxide is present in the solution. Using Beer's law and ultraviolet-visible (UV-Vis) absorption spectroscopy to monitor changes in absorbance elicited by NBT, concentration of superoxide can be recorded and the rate of superoxide formation can be determined (*6*). These colorimetric techniques can determine the effect of the gas environment (inert vs CH_4) and catalyst's presence (with Rh-cat vs without Rh-cat) on $O_2^{\cdot-}$ formation. Moreover, testing the reactivity of synthesized reaction intermediates (CH3-Rh-cat) with commercially purchased samples of the potential oxidants $O_2^{\cdot-}$ and ·OH under non-electrochemical conditions can offer insight into the terminal oxidant. Identifying the ROS and its reactivity can deepen researchers' understanding of compartmentalized ORRs and contribute to optimizing the methane-to-methanol conversion.

MATERIALS AND METHODS

Chemicals and Methods

Commercial reagents for all experimental procedures were purchased from Sigma Aldrich, VWR, and Fisher Chemicals, unless specified. The deionized (DI) water came from a Millipore Milli-Q Water Purification System. The protonated tetra mesityl porphyrin ligand, $(TMP)H_2$, was purchased from Frontier Scientific. NBT and 5,5-Dimethyl-1-pyrroline N-oxide (DMPO) were purchased from Cayman Chemical. 1,2-difluorobenzene (1,2-DFB) was purchased from Oakwood Chemical. If used outside the glovebox, 1,2-DFB was distilled from CaH2 and freeze-pump-thawed or purified with activated 4 Å molecular sieves to remove residual moisture and methanol (CH3OH) impurities respectively. Tetrabutylammonium perchlorate ($TBAClO_4$), purified by recrystallization in ethanol (C_2H_5OH), was used as the electrolyte in all electrochemical measurements.

Ultraviolet-visible (UV-Vis) absorption spectroscopy was conducted on an Agilent Technologies Cary 60 spectrometer. One-dimensional proton nuclear magnetic resonance (^1H NMR) was recorded on a Bruker AV300 (300 MHz) spectrometer and chemical shifts were measured in parts per million (ppm). Deuterated benzene (C_6D_6) from Cambridge Isotope Laboratories was used as the calibrating solvent (128.0 ppm). A gas chromatograph equipped with a mass spectrometer (GC-MS, Agilent Technologies 5975 with Inert XL Selective Detector) was used for CH_3OH determination and quantification. A Mettler Toledo C_20 Coulometric Karl Fischer (KF) Titrator was utilized for the determination of relative water content in neat 1,2-DFB.

Synthetic procedures

The procedures for the synthesis of Rh-cat and CH_3-Rh-cat follow procedures published by Wayland and were utilized in past work published by the Liu group (*3,8*). The silicon nanowire array applied in the electrolysis experiments was synthesized based on work by Huang and was utilized in past work published by the Liu group (*3,9*).

Fluorescent dye experiments

UV-Vis absorption spectra of NBT and potassium dioxide (KO_2) were carried out in the presence of Rh-cat to reduce background interference of Rh complex. 100 µM NBT, 100 µM Rh-cat, and varying concentrations of KO_2 were mixed in the electrolyte solution containing 0.1 M TBAClO4 in 1,2-DFB before transfer into a quartz cuvette for UV-Vis absorption measurements. UV-Vis absorption measurements were taken at a wavelength of λ = 600 nm to not interfere with the characteristic absorption of Rh-cat. The extinction coefficient of NBT in 1,2-DFB at 600 nm was determined to calculate superoxide ($O_2^{\cdot-}$) concentration. 3 mg of KO_2 and 15 mg of 18-crown-6 ($(CH_2CH_2O)_6$) in a 2:3 equivalent were

dissolved in 20 mL of the electrolyte solution. This solution was diluted 4-fold and stored under argon (Ar) at −30º C until ready to use. Once ready, 500 µL of the KO_2 solution was mixed with varying concentrations of NBT and diluted with the electrolyte solution to a final volume of 2 mL. The reaction between KO_2 and NBT ran 8 hours until completion and UV-Vis absorption spectra were taken after.

All electrochemical tests were performed using a Gamry Instruments Interface 1000-E potentiostat. In an Ar-filled glovebox, cyclic voltammetry (CV) and bulk electrolysis experiments were performed with a 500 µM NBT solution in a glass vial, which was equipped with a Teflon top that contained electrode and gas inlet/outlet ports. The sampling electrolysis experiments were conducted in a customized electrochemical reactor with gas inlet/outlet ports (Figure 2). The cell contained 0.1 M TBAClO4 in 1,2-DFB with a Pt wire counter electrode, an Ag^+/Ag reference electrode with a glass frit, and a 3 mm diameter glassy carbon working electrode. Inert gas flow was used as convection and applied potential (E_{appl} = -1.4V) was held constant in the electrolysis experiments. In a typical electrolysis experiment, the electrochemical cell contained 100 µM NBT dissolved in the electrolyte solution, with a Pt wire as the counter electrode and an Ag^+/Ag pseudo-reference electrode with a glass frit. The working electrode was a Si nanowire array with wire length of 15 µm and diameter of 50 nm. A mixture of CH_4 (supplied by Airgas, 99.5%) and house air was introduced into the reactor at a fixed ratio (P_{CH4}/P_{air} = 35) under a constant flow rate of 3.5 standard cubic centimeter per minute (sccm) of methane and 0.1 sccm of air

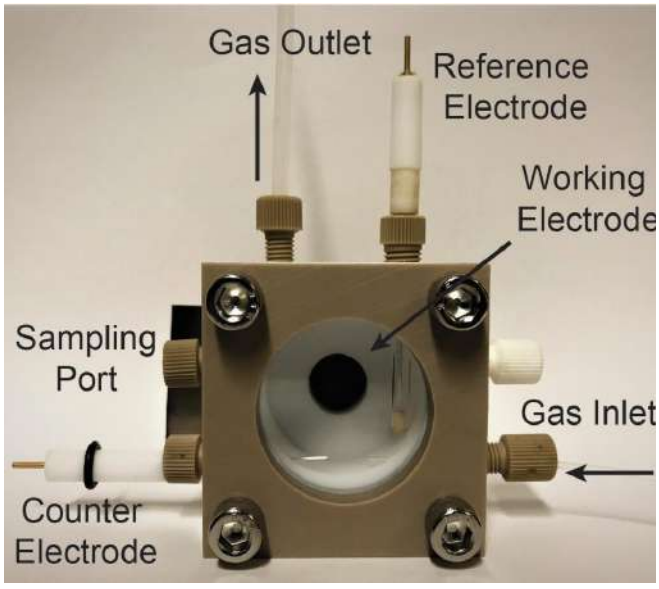


Figure 2. Custom electrochemical reactor used for electrolysis. Electrolysis experiments were conducted in this customized electrochemical reactor. Flow gas (CH_4 or N_2) was pumped into the gas inlet valve using a mass flow controller. The gas outlet valve allowed interaction with the air and prevented pressure build up within the reactor vessel. The counter electrode was a Pt wire, the reference electrode was an Ag^+/Ag wire electrode with a glass frit and the working electrode was a Si nanowire array with wire length of 15 µm and diameter of 50 nm. The sampling port was used to remove the solution for analysis.

with the use of mass flow controllers (Omega Engineering, Inc., Model: FMA5502A 0-10 sccm). The bulk electrolysis was conducted under a constant E_{appl}. Cyclic voltammograms of decamethylferrocene were also conducted to calibrate the potentials of the Ag^+/Ag reference electrode based on the standard potential of decamethylferrocene (-0.059 V vs. Saturated Calomel Electrode, SCE). 200 µL samples were taken every 3 minutes. These extracted samples were diluted 10-fold with the 100 µM NBT solution and their UV-Vis absorption was measured at λ = 600 nm. Additionally, samples were taken before and after the 3 hour bulk electrolysis and were analyzed via GC-MS to observe CH_3OH generation.

Reactivity between CH_3-Rh-cat and ROS

Once the equilibrium $O_2^{\cdot-}$ concentration and total amount of CH3-Rh-cat were quantified from the results of the aforementioned experiments, the ratio between the methylated Rh species and the electrochemically generated $O_2^{\cdot-}$ was postulated. Based on this ratio, 1 mM CH_3-Rh-cat and 20 mM KO_2 were mixed in 1,2-DFB and ^1H NMR spectra were taken every 1 hour. Electrolysis experiments were based on past work by the Lucia and Sullivan group (*10,11*).

Spin trap experiments

DMPO was dissolved and diluted in the electrolyte solution to 200 mM. The DMPO/electrolyte solution was stored at -30º C when it was not in use. For the experiment, 500 µL of the DMPO/electrolyte solution was mixed with KO_2 and cumene hydroperoxide separately until a final DMPO concentration of 50 mM was reached. The samples were then mixed and immediately frozen under liquid nitrogen (N_2) for transport to the EPR facility at the California Institute of Technology. EPR spectra were obtained at room temperature for in situ observation of the electrochemically generated reactive oxygen species. Typical electrolysis experiments were performed off site and transported to the same EPR facility. Once the electrolysis reaction stabilized and $O_2^{\cdot-}$ concentration equibrilized (approximately 30 minutes), DMPO was added into the electrochemical cell to a final concentration of 50 mM. After the addition of DMPO, samples were taken at 5, 15, 30 and 60 minutes and frozen under liquid nitrogen (N_2). Frozen samples were transported to the EPR facility under N2 and EPR spectra were taken at room temperature. The thawed reaction mixtures were left in room temperature conditions (25º C) for 90 minutes before spectra were taken. A qualitative comparison of the spin trapped splitting patterns and relative peak assignments was deemed sufficient to determine the primary electrochemically-generated ROS.

RESULTS

EPR peak assignments and splitting patterns between standard samples of KO_2 and $PhC(CH_3)_2OOH$ were compared with electrolysis experiments with and without the presence of Rh-cat. Spectra comparisons show more qualitative similarities between KO_2 and electrolysis experiments with and without Rh-cat than $PhC(CH_3)_2OOH$ (Figure 3).

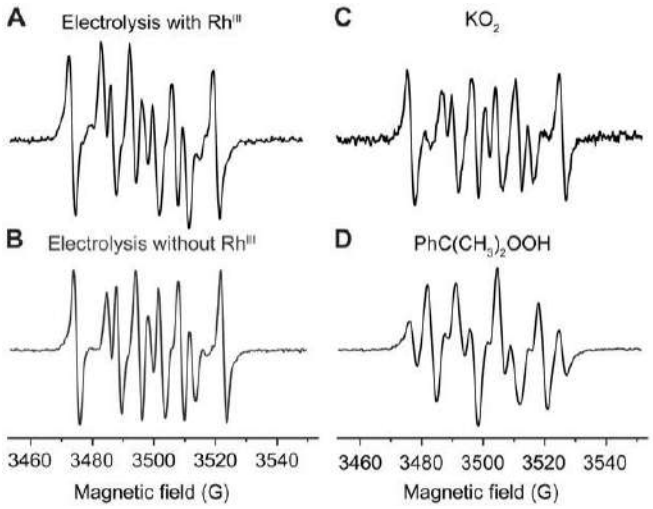


Figure 3. Electron Paramagnetic Resonance (EPR) identification of superoxide as the primary generated oxygen species in electrolysis experiments. To identify the reactive oxygen species (ROS), EPR spectroscopy was conducted with the addition of a spin trap agent, 5,5-dimethyl-1-pyrroline N-oxide (DMPO). Typical electrolysis experiments were conducted and subsequently frozen with liquid nitrogen to transport to the off-site EPR facility at CalTech. DMPO was added to the samples and EPR was run on the samples at varying times after DMPO addition (5, 15 and 60 min) (1). **(A)** and **(B)** indicate typical electrolysis experiments conducted with and without the presence of the rhodium catalyst respectively. **(C)** and **(D)** represent standard samples of potassium superoxide (KO2) and cumene hydroperoxide ($PhC(CH_3)_2OOH$) used for comparison to identify the ROS in electrolysis experiments. The y-axis indicates the relative magnetic spin of a sample at a certain magnetic field strength (G) between 3450G and 3550G. Different paramagnetic species have unique EPR spectra that can be qualitatively identified in solution mixtures. This figure was taken from Figure 2 of reference 4 with permission from the authors (4).

The absence of redox peaks on a cyclic voltammogram of NBT illustrates that NBT undergoes no reaction under the electrolysis conditions used in electrochemical experiments; therefore, NBT can be introduced directly into solution (Figure 4A). The extinction coefficient was calculated from the standard $O_2^{\cdot-}$-NBT curve in Figure 4B to construct the graphs in Figures 4C and 4D, which display the rate of $O_2^{\cdot-}$ formation normalized to the average current response.

CH_3-Rh-cat was also reacted separately under non-electrochemical conditions with KO_2, $PhC(CH_3)_2OOH$, and $(CH_3)_3COOH$. Reactions between CH3-Rh-cat and both PhC(CH3)2OOH and $(CH_3)_3COOH$ generated CH_3OH as a product. However, no CH_3OH was observed when CH_3-Rh-cat and KO_2 were reacted.

DISCUSSION

ROS are regularly observed in cellular respiration and metabolism, so identifying artificial pathways that use oxygen is integral to developing O_2-regulated ambient syntheses (5). The goal of these electrolysis experiments was to identify the primary ROS and gain insight into the oxidative mechanisms of the CH_4 to CH_3OH conversion. Initial EPR results favored the hypothesis that $O_2^{\cdot-}$ serves as the primary oxidant in electrolysis experiments, however, non-electrochemical experiments suggest otherwise. Reactions between methylated rhodium porphyrin (CH_3-Rh-cat) and commercially purchased samples of $O_2^{\cdot-}$ and ·OH yielded results in favor of ·OH as the terminal oxidant. Given these findings, $O_2^{\cdot-}$ is most likely the primary electrochemically generated oxidant in this electrolysis experiment, but not the terminal oxidant. Considering that residual water is always present in the reactor vessel and is a source of inherent error for this experimental setup, it is likely that $O_2^{\cdot-}$ decomposes into ·OH during this conversion. The high reactivity and instability of $O_2^{\cdot-}$ in ORRs also suggests that other side reactions may be responsible for generating the terminal oxidant instead.

Results of UV-Vis spectroscopy reaffirm this conclusion. The existence of an NBT-induced color change supports previous assumptions that superoxide is generated during electrolysis experiments. Usage of the working gas methane versus the inert gas nitrogen had little effect on superoxide formation, except when the rhodium catalyst was added. In the presence of Rh-cat, the rate of superoxide formation slowed, likely due to coordination of Rh-cat with free O_2. Despite the O_2-gradient, it appears that Rh-cat still reacts with O_2 and thus inhibits $O_2^{\cdot-}$ generation, but has little effect on CH_3OH formation. This conclusion is further supported by the similarity in steady-state $O_2^{\cdot-}$ concentrations; the maximum $O_2^{\cdot-}$ concentration generated by the SiNW electrodes is quantitatively similar across every gas and catalyst environment examined in this study. The impact of the catalyst on steady-state $O_2^{\cdot-}$ concentration in a compartmentalized system reinforces the importance of electrochemically-generated O_2 gradients in spatially isolating Rh-cat from O_2 (*3*).

·OH appears to be a terminal oxidant based on experimental observations and known interactions with Rh-C bonds (*3*). Reactions of commercially available .OH species such as cumene hydroperoxide ($PhC(CH_3)_2OOH$) and t-Butyl hydroperoxide ($(CH_3)_3COOH$) with CH_3-Rh-cat successfully led to CH_3OH formation, whereas reactions with KO_2 as the terminal oxidant yielded no CH_3OH. These observations indicate that although $O_2^{\cdot-}$ is electrochemically generated by the SiNW electrodes, it is converted to ·OH prior to CH_3OH formation, likely through protonation in water (*6*).

	$O_2^{\cdot-}$ Rate of formation ($\mu M \cdot mA^{-1} \cdot min^{-1}$)	$O_2^{\cdot-}$ Steady-State Concentration ($\mu M \cdot mA^{-1}$)
$N_2 + Rh^{III}$	16±4	145±27
N_2+ No Rh^{III}	23±3	139±19
CH_4+ Rh^{III}	18±4	128±23
CH_4+ No Rh^{III}	22±4	119±17

Table 1. Superoxide rate of formation and steady-state concentration. Four trials were conducted in the CH4 + air atmosphere with rhodium catalyst (Rh-cat), three trials were conducted in the CH4 + air atmosphere without Rh-cat, and five trials were conducted in the N2 + air atmosphere with and without Rh-cat. Standard deviation was used to construct error bars for all aforementioned experiments. The results were converted to [O2.-] in μM, and normalized to the average current response (in mA) to account for inherent error in the experimental setup. The left column of the table indicates the respective rates of O2.- formation given changes in gas and catalytic environments. The right column compares the average of all steady-state concentrations of O2.- under the same varied parameters. Steady-state was assumed to occur after [O2.-] plateaued.

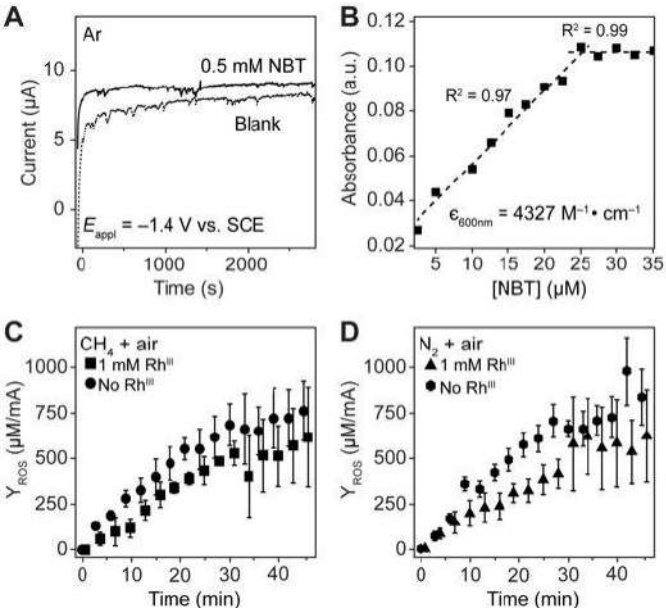


Figure 4. Ultraviolet visible light (UV-Vis) spectroscopic determination of superoxide concentration and rate of formation. (A) illustrates that applying voltage to Nitroblue tetrazolium (NBT) yields no electrochemical response given the absence of redox peaks. **(B)** was used to determine the extinction coefficient using Beer's Law (ϵ600nm= 4327 $M^{-1} * cm^{-1}$). At 25μM, the slope levels off due to stoichiometric interactions between NBT and superoxide (O_2^{-}). Known concentrations of potassium superoxide (KO_2) were monitored with the successive addition of NBT. To quantify O_2^{-} concentration and contribution, typical electrolysis experiments were conducted with the addition of 100 μM NBT, as displayed in **(C)** and **(D)**. Samples were taken every 3 minutes for 45 minutes and their UV-Vis absorption at 600 nm was tracked. Four trials were conducted in the CH4 + air atmosphere with rhodium catalyst (Rh-cat), three trials were conducted in the CH_4 + air atmosphere without Rh-cat, and five trials were conducted in the N2 + air atmosphere with and without Rh-cat. Standard deviation was used to construct error bars for all aforementioned experiments. The results were converted to [O_2^{-}] in μM and normalized to the average current response (in mA) to account for inherent error in the experimental setup. **(C)** and **(D)** show that the gas environment has little effect on the rate of O_2^{-} formation but the presence of a rhodium catalyst can inhibit the rate of O_2^{-} formation. This figure was taken from Figure 3 of reference 4 with permission from the authors (4).

Understanding the mechanism behind this conversion can improve catalyst design and generate environments that better sequester incompatible components of C-H activation and O_2 coordination in the presence of Rh-cat. While silicon is a fairly abundant resource, rhodium porphyrin catalysts are expensive, so more abundant or accessible transition metal catalysts should be investigated as potential replacements. Moving forward, a photoelectrochemical pathway with SiNWs as a photoelectrode can be tested. Such a process would require less applied voltage and would utilize sunlight as an additional energy source for large-scale conversions. This study would investigate the efficacy of SiNWs in large scale electrolysis and the compatibility of this electrode apparatus with photoelectrochemical energy sources. In addition, other natural gas products like ethane (C_2H_6) and propane (C_3H_8) should be tested to determine if the conversion operates similarly on longer alkane chains that contain relatively weaker C-H bonds that are more facile to activate.

ACKNOWLEDGMENTS

The authors report no conflicts of interest and would like to acknowledge Dr. Paul Ayala and Dr. Tama Hasson for their contributions to this work. This work was supported by NIH/NIGMS Grant GM055052. The author would also like to thank Dr. Julie DeLilly-Dumas and Evan Dumas for their endless support.

REFERENCES

1. G. A. Olah, A. Goeppert, G. K. S. Prakash, Beyond Oil and Gas: The Methanol Economy. (Wiley-VCH, 2009).
2. R. K. Srivastava, P.K. Sarangi, L. Bhatia, A.K. Singh, K. P. Shadangi, Conversion of methane to methanol: technologies and future challenges. *Biomass Conv. Bioref.* (2021). doi: 10.1007/s13399-021-01872-5.
3. B. S. Natinsky, S. Lu, E. D. Copeland, J. C. Quintana, C. Liu, Solution Catalytic Cycle of Incompatible Steps for Ambient Air Oxidation of Methane to Methanol. ACS Cent. *Sci.* **5** (9), 1584–1590 (2019). doi: 10.1021/acscentsci.9b00625.
4. B. S. Natinsky, B. J. Jolly, D. M. Dumas, C. Liu, Efficacy analysis of compartmentalization for ambient CH4 activation mediated by a RhII metalloradical in a nanowire array electrode. *Chem. Sci.* **12**, 1818–1825 (2021). doi: 10.1039/D0SC05700B.
5. M. Hayyan, M. A. Hashim. I. M. AlNashef, Superoxide Ion: Generation and Chemical Implications. *Chem. Rev*. **116**, 3029–3085 (2016). doi: 10.1021/acs.chemrev.5b00407.
6. A. B. Grommet, M. Feller, R. Klajn, Chemical reactivity under nanoconfinement. *Nat. Nanotechnol.* **15**, 256–271 (2020). doi: 10.1038/s41565-020-0652-2.
7. K. Makino, T. Hagiwara, A. Murakami, A Mini Review: Fundamental Aspects of Spin Trapping with DMPO. *Radiat. Phys. Chem.* **37** (5), 657–665 (1991). doi: 10.1016/1359-0197(91)90164-W.
8. B. B. Wayland, S. Ba, A. E. Sherry, Activation of methane and toluene by rhodium(II) porphyrin complexes. *J. Am. Chem. Soc.*. **113** (14), 5305–5311 (1991). doi: 10.1021/ja00014a025.
9. Z. Huang, N. Geyer, P. Werner, J. de Boor, U. Gösele, Metal-Assisted Chemical Etching of Silicon: A Review. *Adv. Mater*. **23** (2), 285–308 (2011). doi: 10.1002/adma.201001784.
10. R.-H. Liu, S.-Y. Fu, H.-Y. Zhan, L. A. Lucia, General Spectroscopic Protocol to Obtain the Concentration of the Superoxide Anion Radical. *Ind. Eng. Chem. Res*. **48** (20), 9331–9334 (2009). doi: 10.1021/ie9007826.
11. T. R. O'Toole, J. N. Younathan, B. P. Sullivan, T. J. Meyer, 1,2-Difluorobenzene: a relatively inert and noncoordinating solvent for electrochemical studies on transition-metal complexes. *Inorg. Chem*. **28** (20), 3923–3926 (1989). doi: 10.1021/ic00319a032.

Regional phytoplankton bloom phenologies in the California Current System

Kelsey B. Warren[1], Jordyn E. Moscoso, PhD[2], Colin B. Kremer, PhD[1], Daniele Bianchi, PhD[2]

[1]Department of Ecology and Evolutionary Biology. [2]Department of Atmospheric and Oceanic Sciences.

ABSTRACT

The California Current System (CCS) is a highly productive Eastern Boundary Upwelling Ecosystem that sustains complex and biodiverse coastal ecosystems. Seasonal phytoplankton blooms in the CCS are controlled by wind-driven upwelling of nutrients, influenced by the springtime transition of equator-ward winds. However, variability in topography, stratification, and atmospheric forcing, drive regional differences in bloom phenology. While it is understood that wind-driven upwelling induces phytoplankton blooms, previous studies reported a decoupling between the transition to upwelling-favorable winds and corresponding phytoplankton blooms, making wind a potentially unreliable indicator of upwelling. To provide an alternative indicator of upwelling for comparison to bloom activity, this study used sea surface temperature (SST) to supplement satellite-derived chlorophyll-a concentration (CHL) data in a regional anomaly analysis. This study calculated differences in the timing and strengths of phenological upwelling regimes of four latitudinal regions along the California coast using median climatological values of CHL and SST as well as calculations of the onset dates of blooms and upwelling. The San Francisco (SF) region exhibited the highest intensity of CHL, while the Southern California Bight (SCB) showed the lowest CHL values. Central and Northern California showed the expected spring/early summer blooms, the SCB had the earliest blooms, and SF showed spring and fall blooms. In general, decreases in SST were accompanied by increases in bloom strength throughout the CCS around springtime. These findings provide insight into larger ecosystem changes when shifts in intra-annual bloom timing are observed in the long term and propagate to higher trophic levels.

INTRODUCTION

Eastern Boundary Upwelling Ecosystems (EBUEs) are regions of the global ocean with strong physical and biogeochemical coupling that drive primary productivity and support a robust food web (*1,2*). In EBUEs, upwelling—the process that brings cold, nutrient-rich water to the surface—supports robust primary productivity, or phytoplankton blooms (*3,4,5*). Phytoplankton productivity plays an important role in broader ecosystem dynamics and serves as the base of the food web (*6*).

Understanding bloom activity in EBUEs provides insight into climate and ecosystem changes in EBUEs, such as fish catch, ecosystem composition, ocean hypoxia, and ocean acidification (*7-12*). Furthermore, studying the intra-annual timing of upwelling-driven phytoplankton blooms improves the scientific understanding of patterns of productivity that affect food web composition and structure (*6,13,14*). Specifically, coastal management efforts require a broad knowledge of the seasonal patterns of upwelling to inform policy, identify regions that require intervention, and warn coastal communities about potential impacts to human health (*15*). As such, it is necessary to understand the seasonality and timing of these blooms in the California Current System (CCS) to support management efforts.

Phytoplankton blooms can be remotely sensed with satellite-based measurements of chlorophyll-a concentration (CHL). Because blooms generally result from the increased nutrient content associated with cold, upwelled water, another measurable signal of upwelling-driven blooms is decreased sea surface temperature (SST) (*3*). Figure 1 shows maps of climatological median CHL and SST values from the years 2002 to 2020. Regions of high upwelling generally exhibit colder SST values and higher CHL concentrations. A correlation of climatological nearshore (defined as within 50 km of the coast) SST and log10 (CHL) is found in Section 3.2.3; these results show an expected negative trend between the two variables. The physical forcing that drives upwelling in EBUEs occurs on a seasonal cycle with the intensification of equator-ward winds during the spring (*4,16*).

On longer time scales, nearshore CHL is influenced by basin-scale patterns of temperature including the El-Nino Southern Oscillation, the Pacific Decadal Oscillation, and the North Pacific Gyre Oscillation (*17,18*). These climatic variations are particularly important in the southern latitudes of the CCS, where the nearshore circulation varies from canonical upwelling patterns (*19*). It is unclear, however, whether signals from these basin-scale oscillations are captured in short datasets

over several decades. Likely, more work needs to be done to determine the impact of larger-scale oscillations on phenological measures of upwelling over a longer timescale. As a consequence, this study focuses on regional variability in the timing of upwelling.

Regional variability in upwelling and the resulting phytoplankton blooms are determined by a wide variety of physical parameters. Characteristics such as land topography (*20*), ocean bathymetry (*21*), eddies (*22*,*23*), stratification (*24*), and strength of wind-forcing (*25*) modulate the strength of upwelling (*26*,*5*) and the resulting phytoplankton blooms. Additionally, many of these parameters differ regionally, which causes variability in regional ecological responses to such parameters (*27–29*).

This study augments previously defined methods to identify the timing of phytoplankton blooms in the CCS and discuss the method's performance with a specific focus on regional evaluation. Typically, phenological studies of phytoplankton bloom behavior have been conducted in regions with strong seasonality (*30*). Prior work used physical indices of upwelling, such as wind measurements, to define upwelling-favorable conditions (*26*,*31*). While the CCS does exhibit strong seasonality of upwelling allowing for some predictability of blooms, observational studies recorded that upwelling does not always occur in the presence of upwelling-favorable winds as a consequence of increased stratification (*32*). Henson and Thomas used changes in alongshore wind stress to investigate bloom phenology and found that upwelling-favorable winds generally coincide with blooms in the CCS; however, this study noted recorded instances when wind and bloom timings do not align and suggested that the use of a different physical indicator of upwelling-derived blooms would be more helpful (*28*).

Because studies have shown discrepancies in the wind-predicted seasonal cycle of upwelling and the resulting phytoplankton blooms, satellite derived measurements of SST and CHL were used to answer the following questions: Can SST be used in combination with CHL to describe seasonal patterns of upwelling? What is the timing of the seasonal blooms? How does the timing of phytoplankton productivity vary regionally?

This study specifically investigated the phenological behavior, or seasonal timing, of upwelling in the CCS to establish a climatological understanding of phytoplankton blooms using satellite-derived measurements of SST and CHL. Past studies investigating regional phytoplankton productivity use wind data as a proxy for upwelling; however, this upwelling metric produced inconsistencies with patterns in the measured blooms intensities. To address these issues, this study uses SST as an indicator of upwelling. This choice is justified based on SST being a measure of the physical state of upwelling and not a measure of the forcing that causes upwelling. This study finds that the strongest seasonal cycles of blooms occur in the San Francisco (SF) region followed by the Central California (CC) and Northern California (NC) regions, whereas the Southern California Bight (SCB) shows earlier and less intense bloom occurrences in the calendar year than the other regions. CHL varies greatly from nearshore to offshore (defined as more than 50 km from the coast) longitudes. Finally, it was noted that decreases in SST were accompanied by increases in CHL particularly in the early parts of the calendar year throughout the CCS.

MATERIALS AND METHODS

In this section, the details of the datasets and methods used to understand bloom phenology in the CCS are described.

2.1 Datasets

This study used a merged, daily, satellite-derived CHL dataset to determine the timing of upwelling in the CCS at 1 km resolution. A description of merging the high-resolution, 1 km, surface CHL data is outlined in Kahru et al. using methods from Maritorena and Siegel, with data from California Merged Satellite-derived Dataset (*33*,*34*,*35*). This dataset is tuned to in-situ observations to increase accuracy in surface CHL measurements in the CCS (*36*). This dataset is referred to as SPG-SAT throughout this study.

Sea surface temperature (SST) is Group for High Resolution Sea Surface Temperature (GSRSST) Level 4 MUR Global Foundation Sea Surface Temperature, downloaded from the Physical Oceanography Distributed Active Archive Center (PO.DAAC) (*37*). For use with the SPG-SAT dataset in Section 3.2, this study uses a linear interpolation from a Cartesian mapping to an Albers projection conical grid provided with the 1 km CHL dataset.

2.2 Climatological measures of upwelling

Climatologies for both datasets are calculated from 2002-2020 for days 1-365. Climatologies are calculated using daily averages by Julian day of the year. For example, data is averaged from day 1 (January 1) across the nineteen years, and the average is set as the climatological data value for day 1. Data for leap years is calculated by averaging days 365 and 366 (*38*). This study defines the CCS over a region of 32° N to 40° N, and -130° E to -117° E.

2.2.1 Phenological measures of upwelling

Phenological measures of upwelling were adapted from metrics developed by Henson and Thomas, Henson et al., Siegel et al., and Henson et al. using daily climatologically averaged time-series (*28*,*30*,*39*,*40*). These methods were originally defined to identify the start of seasonal blooms in the North Atlantic by Henson et al. and Siegel et al., and adapted to understand upwelling timing by Henson and Thomas (*40*,*39*,*28*). The metric for upwelling is 5% above the median climatological CHL value, shown in Section 2 for three days. For brevity, this metric is referred to as the median CHL.

Upwelling onset in SST follows the same metric, but with 5% below the median climatological SST value, because patterns of SST are theorized to be inversely proportional to measures of surface CHL (*41*). Similarly, this metric is referred to as the median SST. Both the median CHL and median temperature from 2002-2020 are shown in Figure 1.

Following García-Reyes and Largier, this study defines the "upwelling season" as April to June, due to expected mean up-

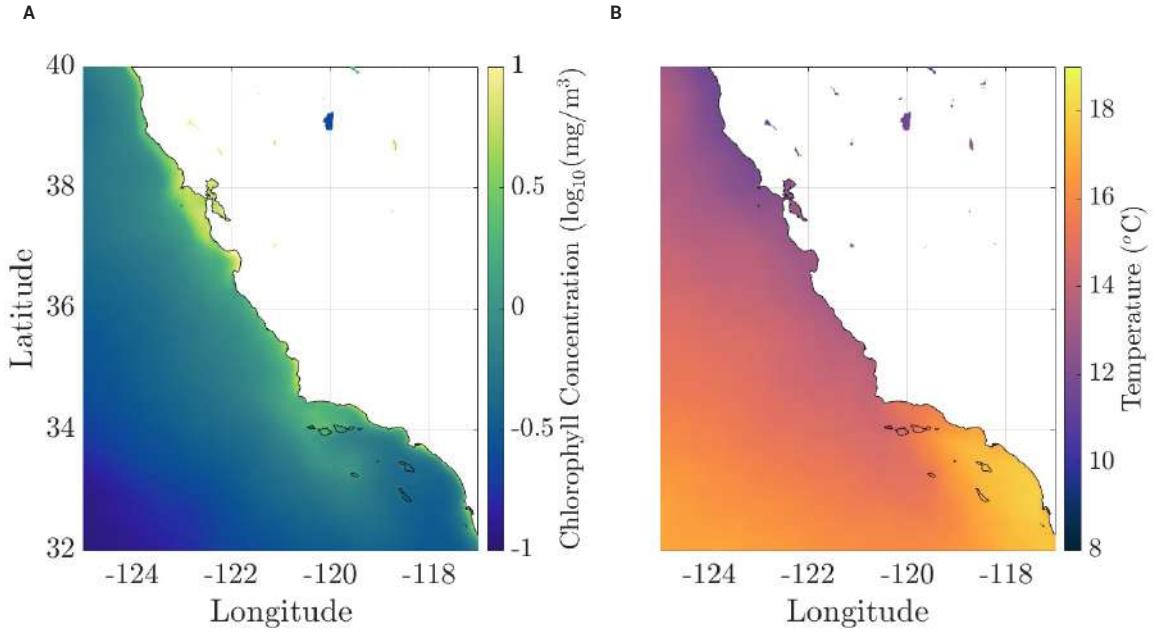


Figure 1. Maps of median chlorophyll concentration and sea surface temperature for 2002-2020 climatology. Maps of median values of chlorophyll (CHL) concentration log10(mg/m3) for 2002-2020 climatology in the California Current System showing **A)** median CHL concentrations from daily, 1 km SPG-SAT merged data (*34*) and **B)** shows median 1 km SST values. Chlorophyll concentration and sea surface temperature should show a negative correlation.

welling-favorable wind forcing (*38*). The study then defines a "relaxation season" between July and September in which upwelling-favorable winds weaken and the ocean re-stratifies. Finally, the "storm season" is defined as December to February, when ocean mixing is driven by atmospheric storm tracks.

2.3 Correlation analysis

This study performs a standard linear regression analysis of nearshore, climatological SST and $\log_{10}$(CHL) regionally to produce slope and R-squared values for each region. Distributions of phytoplankton in EBUEs are often presented as a logarithmic distribution; therefore, in each region, a correlation analysis was conducted between median SST and $\log_{10}$(median CHL) within 50 km from the coast (*5*). A linear relationship between SST and $\log_{10}$(CHL) implies an exponential relationship between SST and CHL in linear space.

This study compares slope and R-squared values between regions to understand spatial differences in the correlation of SST and $\log_{10}$(CHL).

2.4 Available data

Figure 2 shows the fraction of days in the 2002-2020 climatology with CHL data available in SPG-SAT 1 km data. In most of the SCB region, over 60% of days have data present, and in the CC, SF, and NC regions, 40% to 50% of days have data present. The regional availability of data is highlighted because trends in areas with greater data present are more strongly corroborated and are therefore more likely to be consistent patterns. SST data was excluded from Figure 2 because the SST product merges observations and model estimates, and thus is continuous in time.

RESULTS

In this section, trends in phytoplankton bloom phenological upwelling in the CCS are analyzed. This study focuses on phenological patterns of upwelling in the high-resolution CHL dataset from Kahru et al. with SST-determined measures of upwelling in Section 3.2 (*33*).

3.1 Climatological phenology in the California current system

In Figure 1, median CHL and SST values are mapped in the CCS for the 2002-2020 climatology using SPG-SAT 1 km data (a) and SST 1 km data (b), respectively. CHL values are shown on a log scale. Based on Figure 1, areas with lower values of median SST show higher values of median CHL. This trend is consistent with upwelling-derived phytoplankton blooms. The median, rather than the mean, CHL and SST values based on preliminary analysis (not shown) were included to avoid skewing due to outliers. Median values are used as measures of center, and interquartile ranges are used as measures of variability throughout the study.

The phenological analysis was divided into four regions with distinct upwelling regimes. The NC region (latitudes 38.5° N to 40° N) extends from the California-Oregon border to just north of the San Francisco Bay. This region is characterized by a west-facing coast with lower human population density than the SF and SCB regions. The SF region (latitudes 36.5° N to 38.5° N) spans from Monterey to the northern end of the SF bay, where a larger extent of CHL offshore is observed with the lowest SST values. Next, the CC region (latitudes 34.5° N to 36.5° N) is defined from Point Conception north to Monterey. This region is characterized

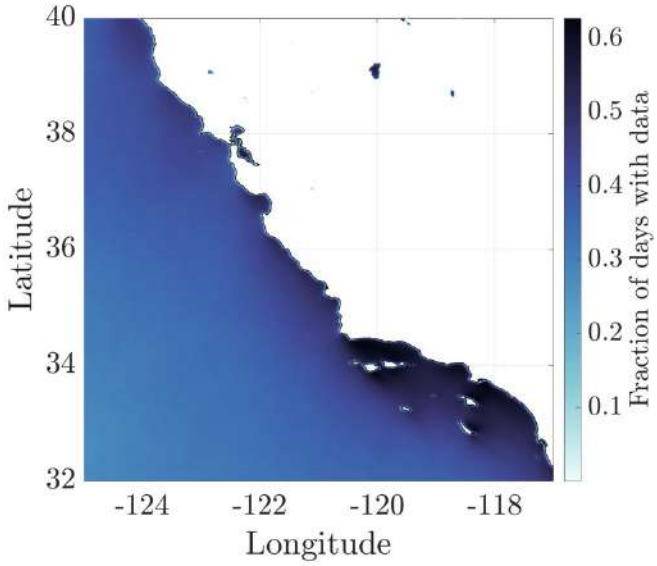


Figure 2. Maps of the fraction of days with chlorophyll concentration data available in 2002-2020 climatology. This figure shows a measure of data available in the 2002-2020 climatology with SPG-SAT 1 km data.

by slightly lower CHL and slightly higher SST than the SF region. In comparison, the SCB region (latitudes 32° N to 34.5° N) has a dynamically different regime with low CHL and high SST values. Past studies have attributed the upwelling behavior in the SCB region to the region's wide and deep continental shelf. A more detailed discussion of these regional qualities is found in Section 3.2.

Comparing measurements of CHL in nearshore and offshore regions, across-shore variability is greater than along-shore variability. 'Across-shore' refers to analysis that focuses on patterns of data perpendicular to the coast, while 'along-shore' describes data parallel to the coast. Nearshore CHL values are much higher than offshore CHL measurements, and concentrations show a decreasing trend moving from nearshore to offshore waters at all latitudes examined in Figure 1. This pattern is consistent throughout the CCS, as the highest CHL values recorded appear closest to the coast, except for elevated CHL levels in the waters surrounding the Channel Islands around 34°N within the SCB region.

Figure 3 shows the median along- and nearshore CHL averaged as a function of latitude per year-day of the 2002-2020 climatology. The green line and shading represent median CHL values and interquartile ranges, respectively, for SPG-SAT 1 km data.

In Figure 3A, an increasing trend in median CHL values beginning around day 70 of the year (March 10) indicates a spring bloom with peak intensities around days 100 and 180 (April 10 and June 29). Median CHL values give different concentrations throughout most of the climatological year, with notable exceptions during which the datasets show similar median concentration values: from approximately days 180 to 225 (June 29 to August 13) and around late December to early January, days 350 to 5 (December 16 to January 5). The summertime period from days 180 to 225, when the median CHL values between datasets align, coincides with a sustained dip in CHL during which no major blooms are detected. Another bloom during fall occurs around days 250 to 310 (September 7 to November 6) of the climatological year.

Not only are seasonal changes in the intensity of CHL seen in Figure 3, but temporal differences in the sizes of the interquartile ranges of the CHL data are also present. In periods of peak blooms, there is much more variability in CHL, as seen in the wider interquartile range shading in Figure 3 during blooms. The width of the interquartile range increases beginning around day 60 (March 1) when CHL begins to increase at the start of the spring bloom. Times of the year with wider interquartile ranges result from greater latitudinal variability in CHL.

Figure 3B shows the temporal trends in median SST values for the 2002-2020 climatological year. The slight decreasing trend in SST around days 80 to 100 (March 21 to April 10) could be due to upwelling, in which cold, deep waters rise and decrease the temperature of surface waters. This decrease in SST coincides with the rise of median CHL values shown in Figure 3, beginning around day 70 and peaking around day 100 (March 11 and April 10). SST values increase throughout summertime and reach a maximum around day 250 (September 7), followed by a decreasing trend into late fall and early winter.

3.1.1 Onset of upwelling

Figure 4 shows maps of the date of the onset of upwelling in the CCS for the 2002-2020 climatology calculated using SPG-SAT 1 km CHL data (a) and SST 1 km data (b). The date of onset of upwelling were calculated using these conditions: the day on which CHL first reaches and remains at 5% above the annual median value for three days and the day on which SST first reaches and remains at 5% below the annual median value for three days. The onset date only reflects the first date on which these conditions are met, regardless of if the conditions are met more times throughout the year. This analysis is concerned with the upwelling-associated spring bloom, which tends to be the first bloom in each calendar year.

The map in Figure 4A shows the date on which the conditions for the onset calculation are first met during the 2002-2020 climatological year for CHL. Throughout most of this nearshore region of the CCS, the conditions for CHL-determined onset have already been met on day 1 (January 1) of the climatological year, seen in the dark blue color present throughout most of the map. There is a region of later CHL-determined onset from approximately 36° N to 38° N, off the coasts of San Francisco Bay and Monterey Bay (in the northern end of the CC region and most of the SF region). The spring CHL-determined bloom in this region starts around days 60 to 70 (March 1 to March 11) of the climatological year. There is another region with spring CHL-determined onset from just north of Point Conception around 35°N extending southeast down to about 33°N (in the southern end of the CC region and the northern end of the SCB region). The CHL-determined onset here occurs around days 20 to 30 (January 20 to January 30) of the climatological year. It appears that

the CHL onset conditions have already been met on day 1 (January 1) in the SCB region.

In Figure 4B, the map shows the day on which SST values first meet the conditions for the onset of upwelling. In nearshore waters along the coast of the CCS from 40°N to 34.5°N (the NC, SF, and CC regions), the SST-determined onset is fairly regular around day 50 (February 19). The SST-determined onset is delayed to about day 100 (April 10) in the northern end of the SF region around 38°N and in the southern end of the SF region around 36.7° N (Monterey Bay). An area extended southeast from Point Conception within the SCB region similar to that in the CHL-determined onset maps is present in the SST-determined onset map with onset dates around days 40 to 50 (February 9 to February 19). Conditions for SST-determined onset have been met already by day 1 (January 1) of the climatological year in the rest of the SCB region.

3.2 Regional phenology

3.2.1 Along-shore variations in phenology

Figure 5 shows climatological median CHL values for the SPG-SAT 1 km data as a function of latitude nearshore in the CCS. There are strong peaks in CHL present around 37.5°N (the middle of the SF region) and 34.5°N (the border of the CC and SCB regions), which are located off of San Francisco Bay and Point Conception, respectively. Regions of upwelling experience highly variable CHL throughout the climatological year because of the seasonal nature of the physical forcing that results in strong phytoplankton blooms. CHL was analyzed over SST by latitude because SST increases with decreasing latitude due to increasing solar radiation closer to the equator. This radiation-induced constant rise in SST moving down the California coast limits the ability to discern between location-dependent versus upwelling-driven low SST values.

It is noted that there are variations in the width of the interquartile range of CHL latitudinally. Examining Figure 5 with a focus on the interquartile range, the width of the shading of interquartile ranges increases in areas with stronger upwelling regimes. The widest interquartile ranges are seen between 36.5°N and 38.5°N (the entire SF region) and between 33.5°N and 35.5°N (the northern end of the SCB region and the southern end of the CC region).

Previous studies have pointed to drivers of alongshore variability (*42*). These drivers include variations in along-shore wind-stress, bottom topography, and stratification (*43,21,44*). These physical parameters can account for along-shore variability in the median, however, drivers were not explored in this study; instead, along-shore regimes were defined based on the different behaviors in the median.

3.2.2 Regional measures of upwelling

This section examines regional differences in the phenology of blooms throughout the CCS. This study separates the CCS into four latitudinal nearshore regions informed by the analysis conducted in Section 2 and Figure 5.

Figure 6A shows median, along- and nearshore (50 km) averaged CHL values in the NC region, from latitudes 38.5°N to 40°N. Median CHL values and associated interquartile ranges are plotted per day of the 2002-2020 climatological year. Beginning analysis at the start of the climatological calendar year, a weak bloom is detected around day 80 (April 20), followed by a more intense bloom peaking around day 180 (June 28). CHL values are somewhat elevated around days 250 to 300 (September 6 to October 26) and then decrease toward the end of the year.

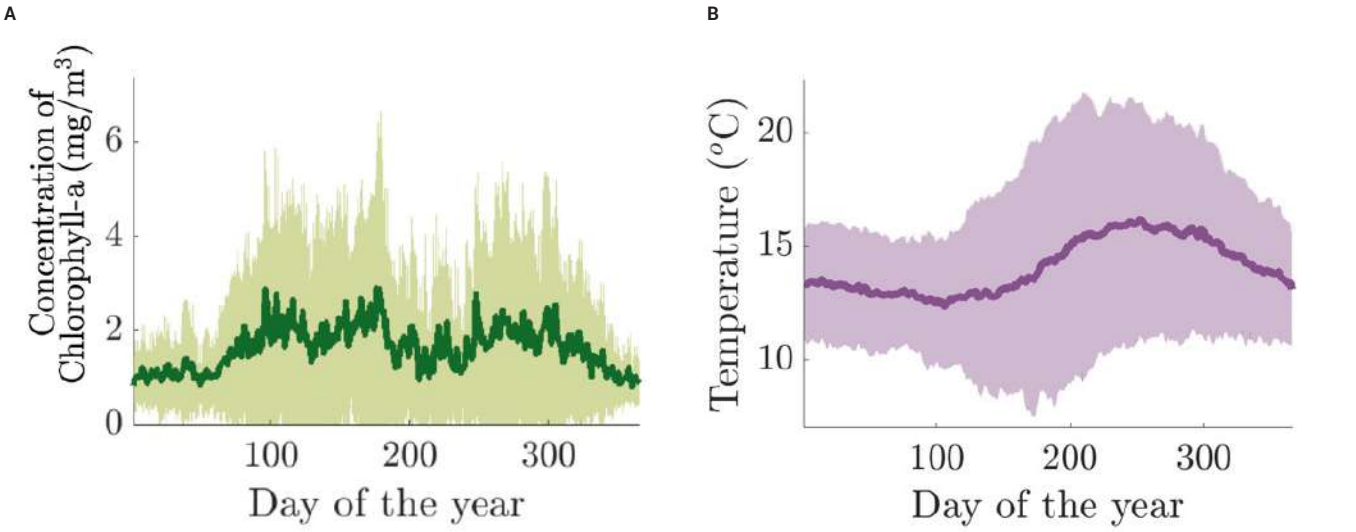


Figure 3. Median chlorophyll concentration and sea surface temperature per Julian day of the 2002-2020 climatology. A) Median, along- and nearshore (within 50 km of the coast) averaged, chlorophyll values are shown per year-day of the 2002-2020 climatology. The line and shading show median chlorophyll-a concentration values and associated interquartile ranges, respectively. **B)** Median, along- and nearshore averaged, sea surface temperature values are shown per year-day of the 2002-2020 climatology. The line and shading show median sea surface temperature values and associated interquartile ranges, respectively.

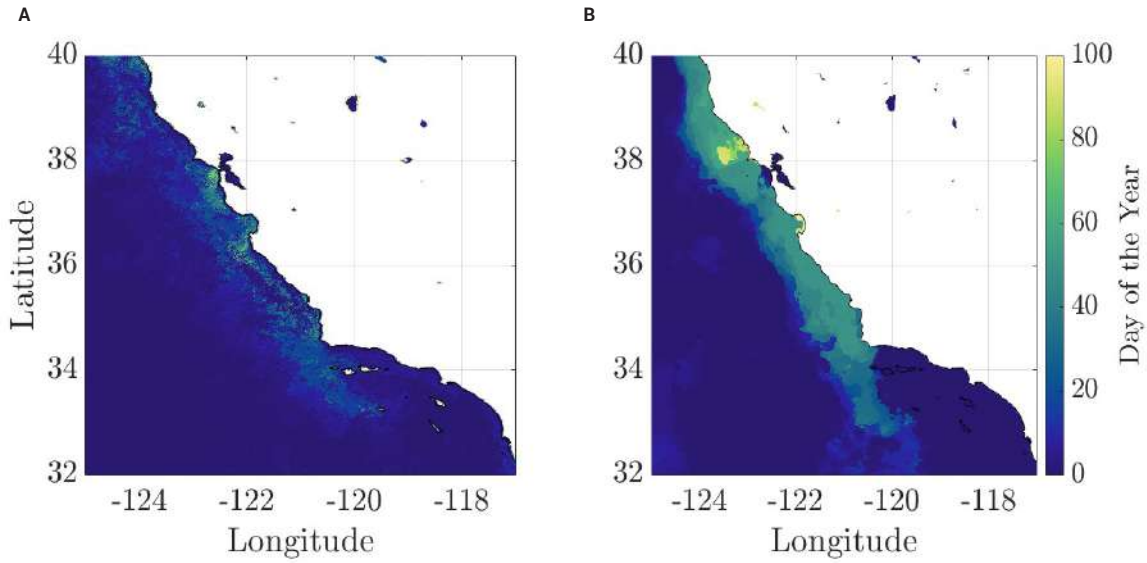


Figure 4. Maps of calculated bloom onset using chlorophyll concentration and sea surface temperature. Maps showing the calculated day of bloom onset with **A)** measured surface chlorophyll concentration and **B)** measured sea surface temperature in the 2002-2020 climatology. The conditions used for the onset date calculation are as follows: the day on which chlorophyll concentrations first reached and remained at 5% above the annual median value for three days, and the day on which sea surface temperatures first reached and remained at 5% above the annual median value for three days. There is consistency in the results between onset date calculated by **A)** chlorophyll concentration and **B)** sea surface temperature, with a nearshore (within 50 km of the coast) band exhibiting delayed onset in comparison to the rest of the California Current System. See Figure 8 for similar analysis conducted using a 5% rather than 10% metric for calculating onset date.

Figure 6B shows median, along- and nearshore (50 km) averaged CHL values in the SF region, from latitudes 36.5°N to 38.5°N. Median CHL values and associated interquartile ranges are plotted per day of the 2002-2020 climatological year. Starting at the beginning of the year, an increase in median CHL starts around day 30, and CHL continues to rise until a peak around day 130 (January 30 and May 10, respectively). A second peak occurs around day 190, and CHL values decrease until about day 225 (July 9 and August 13, respectively) in late summer. A fall bloom is apparent in the 1 km data, with median CHL increasing from day 225 to 275 (August 13 to October 2). After the fall bloom, CHL declines rapidly to day 365 (December 31).

Figure 6C shows median, along- and nearshore (50 km) averaged CHL values in the CC region, from latitudes 34.5°N to 36.5°N, between Monterey Bay and Point Conception. Median CHL values and associated interquartile ranges are plotted per day of the 2002-2020 climatological year. Beginning on day 1 (January 1), CHL begins increasing around day 60 with a spring bloom peak around day 100 (March 1 and April 9, respectively). A second, more concentrated bloom peaks around day 180, after which CHL decreases until a sustained fall bloom occurs between days 250 and 300 (June 29, September 7, and October 27, respectively). CHL decreases from days 300 to 365 (October 27 to December 31).

Figure 6D shows median, along- and nearshore (50 km) averaged CHL values in the SCB region, from latitudes 32°N to 34.5°N. Median CHL values and associated interquartile ranges are plotted per day of the 2002-2020 climatological year. Beginning just after day 1, a CHL bloom appears, peaking between days 5 and 20 (January 1, 5, and 20, respectively). After this winter bloom, CHL remains low until about day 50, when it gradually begins to increase, and day 70, when it rises rapidly (February 19 and March 11, respectively). CHL remains high from days 70 to 120 (April 30). CHL then drops rapidly around day 120 and remains low for the rest of the climatological year with the exception of one peak around day 340 (April 29 and December 6, respectively).

3.2.3 Regional correlation analysis of SST and $\log_{10}$(CHL)

This study evaluated the relationship between median SST and $\log_{10}$(CHL) via linear regression with the intention to understand the correlation between cold-upwelled waters and phytoplankton blooms. Using SST as an indicator of upwelling, a negative correlation between SST and $\log_{10}$(CHL) was expected, which implies that the upwelling of cold, nutrient-rich water supports phytoplankton blooms. Over the entire CCS, there is little correlation between SST and $\log_{10}$(CHL) (figure not included).

The results of latitudinal correlation analysis in Figure 7 are considered next. Figure 7A shows the results of this analysis in the NC region. This region showed the strongest negative correlation between SST and $\log_{10}$(CHL) with a slope of -0.73 and an R-squared value of 0.85. In the NC region, lower SST values are highly correlated to higher $\log_{10}$(CHL) values, and the high R-squared value indicates that there is less variance in this negative relationship in this region. The results of the correlation analysis for the SF region are shown in Figure 7B. The correlation between SST and $\log_{10}$(CHL) is weakest in the SF region with a slope of -0.02 and an R-squared value of 0. This small correlation value implies that there is almost no linear relationship between SST and $\log_{10}$(CHL) in this region, and R-squared is 0 shows that this slope accounts for none of the variance in the data. In the CC region, Figure 7C shows a -0.16 slope and 0.08 R-squared

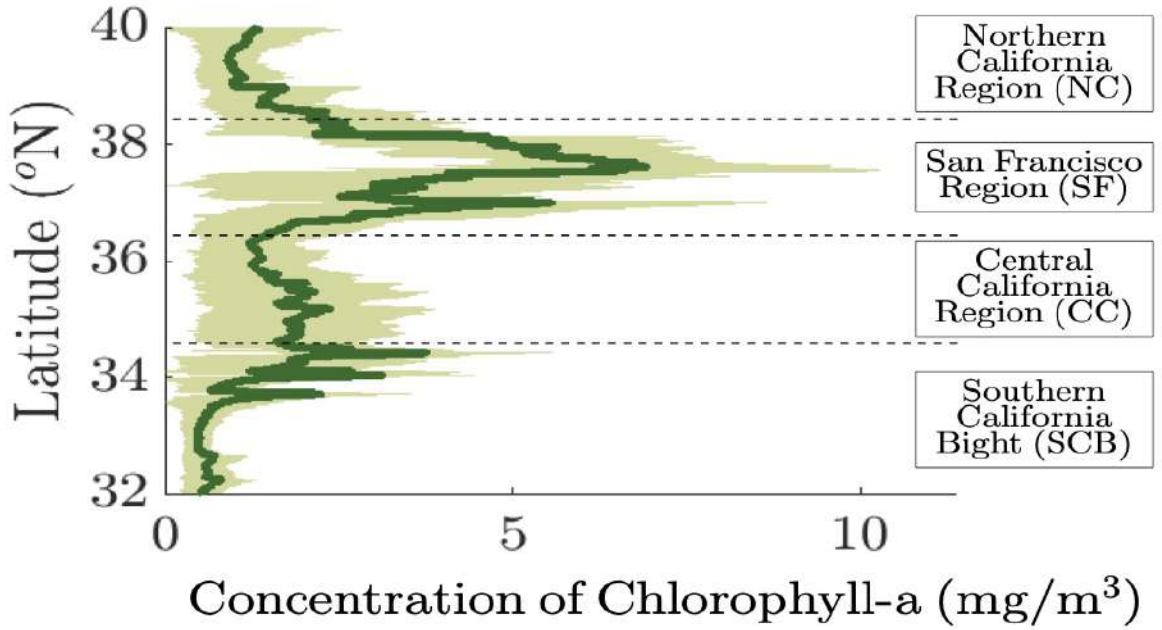


Figure 5. Median chlorophyll concentration by latitude for 2002-2020 climatology. Median, daily-averaged chlorophyll values and associated interquartile ranges are shown across latitudes from 32°N to 40°N for the 2002-2020 climatology in the nearshore 50 km band. The line represents median chlorophyll values, and the shading shows the interquartile range of the chlorophyll values. Four latitudinal regions are separated by horizontal dashed lines with region names labeled in boxes.

value. These values demonstrate a weak negative relationship between SST and log10(CHL) in the CC region. Lastly, Figure 7D shows the correlation for the SCB region. The negative correlation between SST and $\log_{10}$(CHL) in this region is stronger than in the SF and CC regions with a slope of -0.24 and an R-squared value of -0.37. Colder SST values are more correlated with higher $\log_{10}$(CHL) values in the SCB region.

3.2.4 Overview of regional variations in bloom phenology

Overall, the analysis in this study found highly different patterns in phytoplankton growth between the latitudinal regions. Bloom intensities were strongest in the SF region, followed by the CC region, the NC region, and the SCB region. In the NC region, weak CHL signals were observed compared to the other regions. Weak blooms occur in this region in early spring and early fall, with stronger blooms taking place in early to mid summer. In the SF region, strong CHL signals were detected in the climatology in the spring, summer, and fall. In the CC region, blooms were also present in these three seasons, however, the summer bloom was much stronger than the spring and fall blooms in this region. In the SCB region, the strongest blooms occur in late winter to early spring, and there are additional, less intense blooms occurring in December and January. The timings of the blooms in the SCB region are different from those in other regions in the CCS, which could be a consequence of the choice of statistics or other variability in physical forcing.

In the study of the timing of nearshore SST, it was found that patterns in SST data are indicative of seasonal upwelling. The results identified a drop in SST from February to April, as well as an increase in CHL values in both resolutions of data that appears at this time of year.

The same methods were applied to indicate the start of the blooms in the CCS using SST data, which could indicate the physical start of upwelling. In both data resolutions, the results of this analysis show evidence for a lag in bloom onsets off the coasts of the San Francisco and Monterey Bays (in the SF region and the northern end of the CC region), compared to the alongshore averaged start of upwelling. A weaker lag signal was also observed ranging from Point Conception south past the Channel Islands (the northern end of the SCB region). This offshore area south of Point Conception and west of the Channel Islands exhibits patterns in phytoplankton blooms that are more similar to the rest of the nearshore CCS than does the nearshore area of the SCB region. In the SST data, the onset of decreased temperatures was overall consistent in a nearshore band from 40° N extending south from Point Conception to about 33°N (throughout the NC, SF, CC, and northern SCB regions), with the exception of two later onset patches in the SF region adjacent to San Francisco and Monterey Bays. The conditions for the onset of decreased SST have been met by day 1 of the climatological year in the SCB region.

DISCUSSION

This work provides insight into regional phenological patterns in phytoplankton blooms in the CCS using SST as a physical metric of upwelling. This work expands on the work of Henson and Thomas, which found that, over the eight year period from 1998

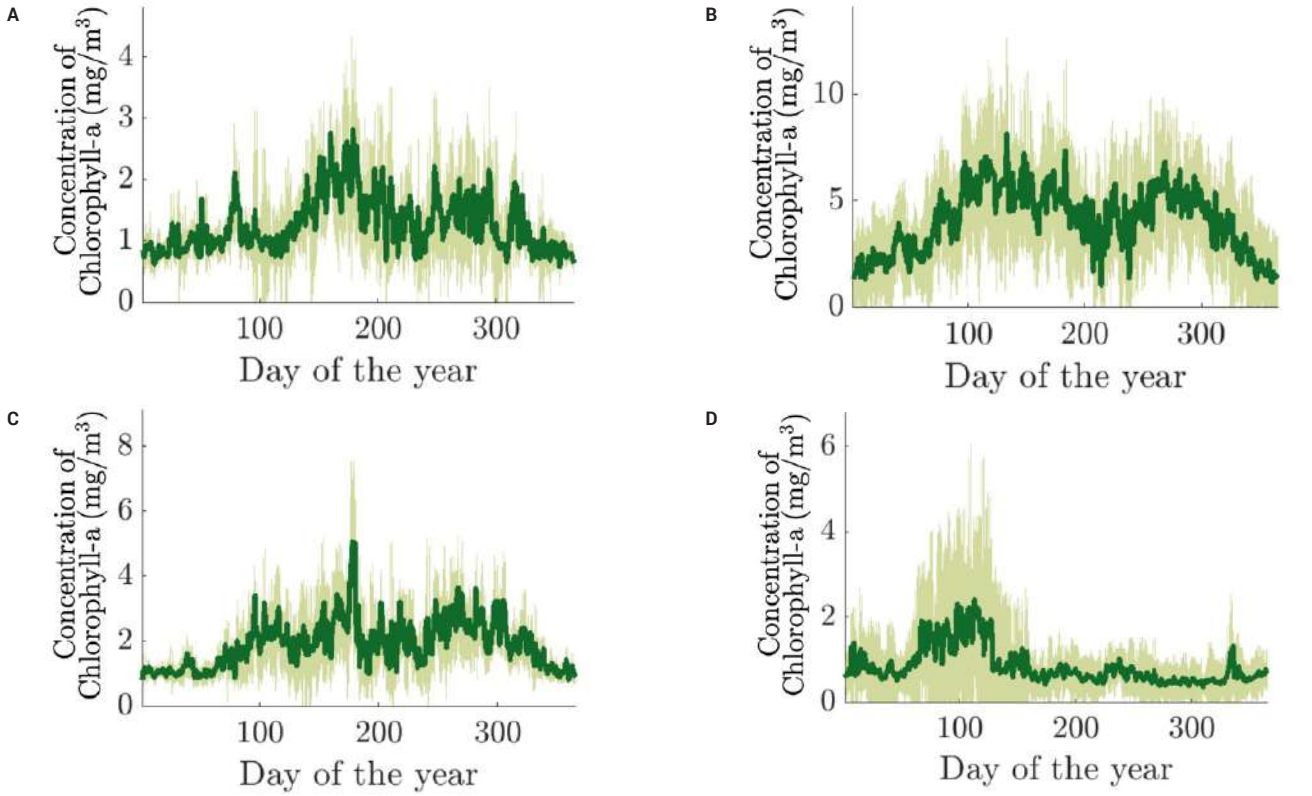


Figure 6. Regional plots of chlorophyll concentration per year-day of 2002-2020 climatology. Median, along- and nearshore (within 50 km of the coast) averaged, chlorophyll values and associated interquartile ranges are shown per year-day of the 2002-2020 climatology for **A)** Northern California, **B)** San Francisco, **C)** Central California, and **D)** Southern California Bight.

to 2005, patterns were easily identifiable in determining bloom timing in highly seasonal regions, but less so in regions with less distinct seasonality of blooms (*28*). The results also align with this finding (Figure 6). This agreement is likely due to the consistency between the frameworks for defining the onset dates of blooms, which was less accurate in the SCB region, which lacks strong seasonality, compared to in the CC, NC, and SF regions. Similar to Henson and Thomas, the clarity of the results was improved by reducing the latitudinal extent of minor regions in the analysis (*28*). The results identified high variation in bloom phenologies and intensities between the NC, SF, CC, and SCB regions (Figure 7). Regional variation was also present in the degree of negative correlation between nearshore SST and $\log_{10}$(CHL). Additionally, this study found strong variation in across-shore measures of CHL, as did Henson and Thomas in their oligotrophic, offshore, and upwelling regions (*28*). One common limitation to the methods and those used in Henson and Thomas is that it can be difficult to distinguish between potential causes of blooms, which may include upwelling, stormwater runoff, and other anthropogenic inputs, which would indeed affect bloom timing in regions with strong anthropogenic forcing (*28,2,3,23,11*).

Previous studies have shown that a wide range of physical parameters impacts phytoplankton blooms, including the width of the continental shelf and shielding caused by local topography, stratification, and wind (*20,21,45,46*). There seems to be a progression of the onset of blooms moving northward in the CCS, with the earliest bloom start dates located in the SCB region and later ones occurring in the SF region, although this pattern has not been studied extensively. Not only is this latitudinal onset pattern present in measurements of CHL, but it also appears in SST data, indicating that this is a robust feature of the data, and likely driven by upwelling. The use of alternative statistical metrics in the calculations could provide an improved understanding of latitudinal patterns in the timing of the onset of spring phytoplankton blooms.

The impacts of this work span a broad range of fields. These include public health precautions for coastal access and recreation, ecosystem and fisheries health, and implications for large-scale ocean trends like ocean acidification, hypoxia, and anthropogenic eutrophication (*47*). A comprehensive understanding of the climatological timing of phytoplankton blooms in the CCS has applications in public health for predicting the occurrence of algal blooms that have the potential to be harmful to human health. Although the causes of harmful algal blooms (HABs) are still not well understood, predicting seasonal patterns of blooms can inform potential patterns of HABs. The results could additionally inform ecologists about the predictability of bloom effects on larger-scale ecosystem dynamics, as changes in phytoplankton growth could exert considerable impacts on other trophic levels through trophic cascades. EBUEs are important fisheries for the world food

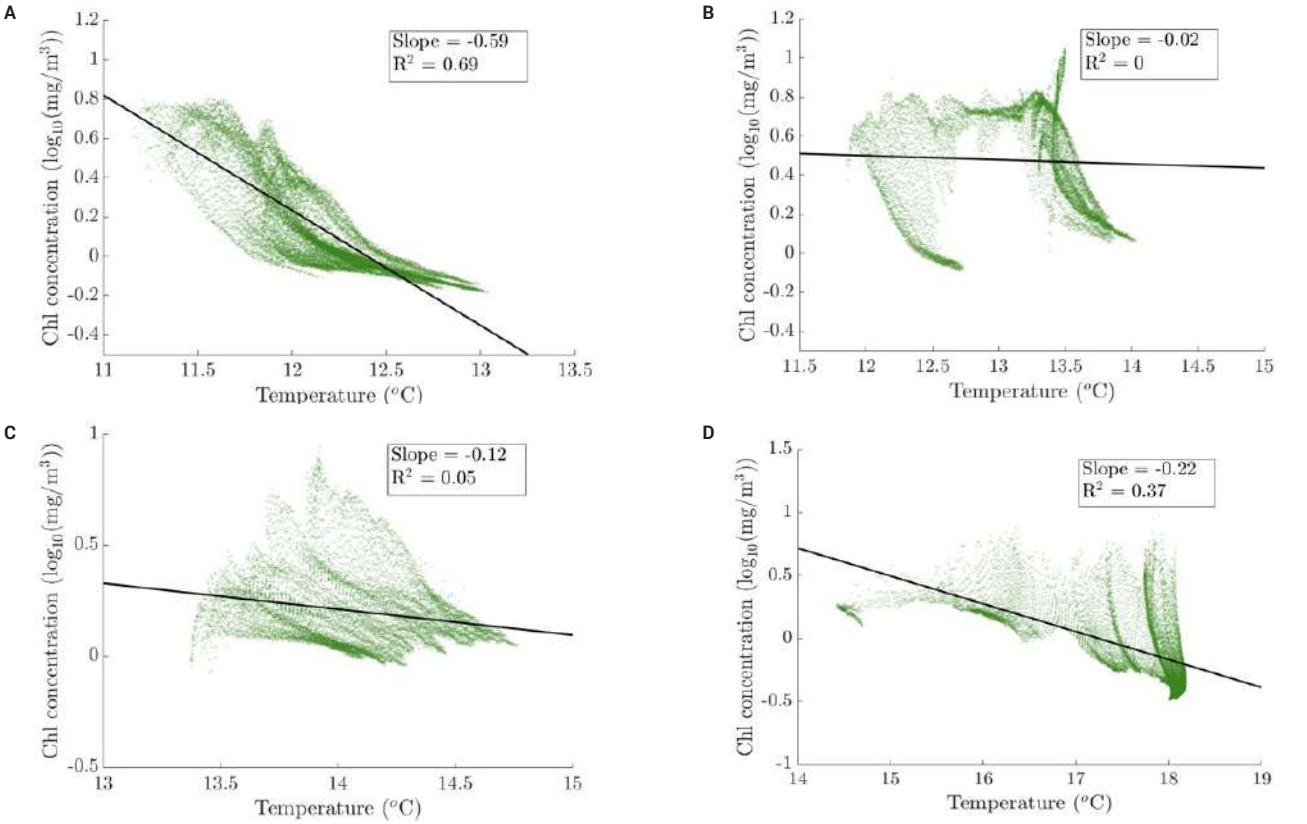


Figure 7. Regional correlations of sea surface temperature and $\log_{10}$(chlorophyll concentration). Correlation of 2002-2020 climatological median values of nearshore sea surface temperature and log10(chlorophyll-a concentration) in **A)** the Northern California, **B)** the San Francisco, **C)** the Central California, and **D)** the Southern California Bight regions. **A)** The Northern California region shows the strongest correlation with the most negative slope and highest R-squared, **B)** The San Francisco and, **C)** Central California regions show very low correlation with slopes and R-squared values near zero, and **D)** the Southern California Bight region exhibits a weak negative correlation with a small negative slope and an intermediate R-squared value.

supply, and understanding the timing of blooms and resulting fish population booms is useful to increase fishing industry efficiency (*8*,*5*). In studying ocean changes as part of global climate change, it is helpful to have a baseline knowledge of the current phenology of blooms for comparison to past and future trends (*7*). Patterns of phytoplankton blooms can be utilized to monitor the effects of climate change processes that exert influence at the planktonic level.

Perhaps the largest limitation to this study had to do with the strict statistical conditions used to identify the blooms. The onset date of phytoplankton blooms was calculated as the date on which CHL first rises to 5% above the annual climatological median and remains above this threshold for three consecutive days. Although the choice of 5% variability was informed by Henson et al., the actual value was chosen independently of the data itself (*40*). Similar choices were used for SST. To test the effects of the choices for statistical thresholds, this study explored other parameters for calculating onset in an effort to better reflect actual bloom phenologies, including using 10% above/below the median (Figure 8) rather than 5%. In the CHL data, as the percentage above the median required for onset was increased, the onset dates nearshore became more delayed in some areas. Increasing this percentage threshold in the SST data results in delayed onset dates of the SST-drop in the majority of the nearshore region, most notably in the SCB region. These choices of thresholds would likely help indicate strong blooms, however more sensitive statistics could be used to indicate weak phytoplankton blooms.

Another limitation to this study was the choice of the date on which the analysis began the phenological search. Because blooms are detected in the climatological wintertime in the SCB region, beginning the search for the onset date around mid-spring (day 50) could result in onset maps that better reflect the start of the late springtime bloom. Alternatively, starting the search for the onset date in fall could improve the analysis, as during this period SST begins to decrease and blooms often occur. This analysis could also use a longer time criterion than three days as the number of consecutive days that CHL must remain elevated in order to define a more sustained bloom. Although the analysis could have better represented the phenology of blooms and drops in SST using different statistical values in the onset thresholds, the results using these metrics are meaningful in that they provide a statistically consistent update to the work of Henson and Thomas (*28*).

While these statistics are used to indicate blooms in CHL, different metrics may be better suited to indicate the start of upwelling in an SST product. For example, a possible metric could

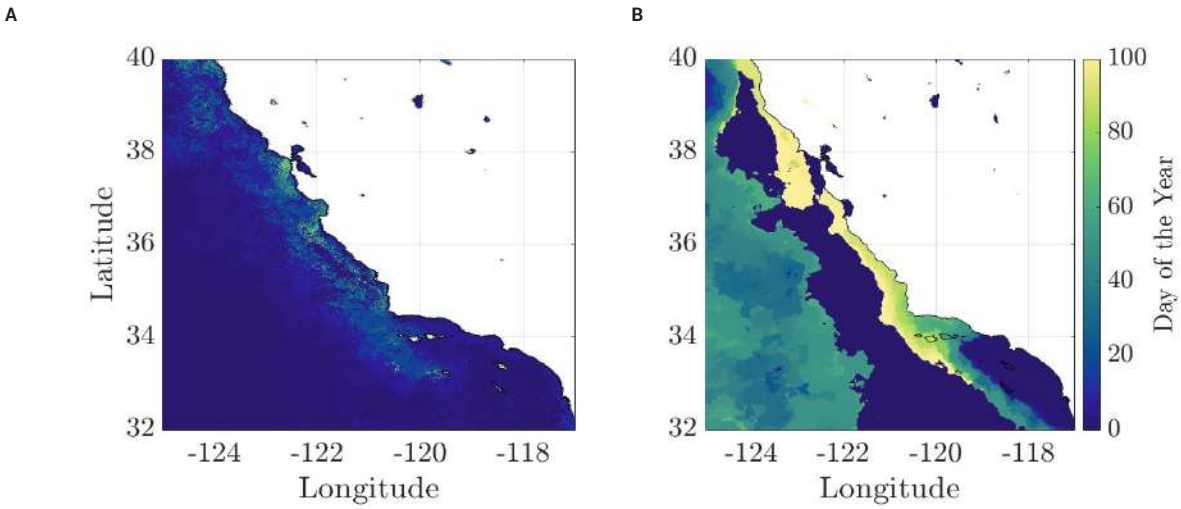


Figure 8. Maps of calculated bloom onset using chlorophyll concentration and sea surface temperature. Maps showing the calculated day of onset of blooms with **A)** measured surface chlorophyll concentration and **B)** measured sea surface temperature in the 2002-2020 climatology. The conditions used for the onset date calculation are as follows: the day on which chlorophyll concentrations first reached and remained at 10% above the annual median value for three days, and the day on which sea surface temperatures first reached and remained at 10% above the annual median value for three days. **B)** shows a delayed onset of upwelling in terms of sea surface temperature in comparison to **A)** an earlier onset of upwelling based on chlorophyll concentration. maybe comment on the visual trends that are important to the paper's findings. See Figure 4 for similar analysis conducted using a 5% rather than 10% metric for calculating onset date.

be deviations from a moving average in SST over shorter time frames. Because SST is often correlated with seasonal solar insolation, deviations from a moving average could serve as a more sensitive indicator of upwelling events. Furthermore, the El Niño Southern Oscillation (ENSO) and Pacific Decadal Oscillation (PDO) strongly impact SST, specifically in the northern CCS. To improve upon the agreement between the CHL and SST onset maps, the analyses of median CHL and SST can be used to inform the understanding of patterns in timing between the two variables (Figures 3, 5, 7). Optimizing the statistical parameters for calculating onset dates can improve temporal association between CHL and SST onsets. The study may be further benefited by conducting spatio-temporal averaging of the datasets.

Future work in understanding bloom phenology in the CCS would certainly benefit from the inclusion of additional datasets such as riverine runoff and wastewater treatment plant sewage outflows, as these could be useful in understanding the impact of anthropogenic input on blooms. It would also be helpful to utilize data on the community compositions of phytoplankton in the CCS to inform the understanding of the roles of certain taxa in bloom timing, strength, location, etc. Examples of these datasets include the Southern California Coastal Ocean Observing System's riverine runoff data and CalCOFI's phytoplankton species counts (*48,49*). The use of these datasets could clarify a drawback of the analysis by distinguishing between upwelling and various anthropogenic inputs as the ultimate drivers of nearshore blooms.

This work investigates the use of SST as a marker of upwelling-induced phytoplankton blooms and finds regional variations in bloom timing and strength throughout the CCS. The design of the study expands on previous work by Henson and Thomas on regional phenological bloom variations in the CCS using SST as an indicator of upwelling events rather than the wind indices employed by Henson and Thomas (*28*). Future work in the field could analyze SST data with different statistical metrics and take into account ENSO and PDO in understanding bloom timing. Additional datasets including anthropogenic input and community composition could also augment studies of phytoplankton phenology. Studies on this topic have implications for the fields of public health, ecology, fisheries, and climate science.

ACKNOWLEDGMENTS

The author would like to thank all the members of the Kremer Lab for their continual support and consistent input of new ideas, as well as Nicole Crane for her meaningful guidance and the connections she provided. The author would also like to thank Morgan McMasters and Truc Pham for conducting the preliminary data analysis which formed the basis of this work. This work was supported by the National Science Foundation, the NOAA ROSES Physical Oceanography program, and the Alfred P. Sloan Foundation.

REFERENCES

1. A. Bakun, R. H. Parrish, Turbulence, transport and pelagic fish in California and Peru current systems. *Calif. Coop. Oceanic Fish. Invest. Rep.* **23**, 99–112 (1982).
2. M. Messié et al., Potential new production estimates in four eastern boundary upwelling ecosystems. *Prog. Oceanogr.* **83**, 151–158 (2009).
3. M. R. Abbott, P. M. Zion, Satellite observations of phytoplankton variability during an upwelling event. *Cont. Shelf Res.* **4, no. 6,** 661–680 (1985).
4. W. S. Wooster, A. Bakun, D. R. McLain, Seasonal upwelling cycle along the eastern boundary of the North Atlantic. J. Mar. Res. **34, no. 2**, 131–141 (1976).
5. F. P. Chavez, M. Messié, A comparison of Eastern Boundary Up-

welling Ecosystems. *Prog. Oceanogr.* **83**, 80–96 (2009).
6. C. A. Stock et al., Reconciling fisheries catch and ocean productivity. *P. Natl. Acad. Sci. USA*. **114**, E1441–E1449 (2017).
7. R. R. Rykaczewski, J. P. Dunne, Enhanced nutrient supply to the California Current Ecosystem with global warming and increased stratification in an earth system model. *Geophys. Res. Lett*. **37** (2010).
8. J. E. Moscoso, D. Bianchi, A. L. Stewart, Controls and characteristics of biomass quantization in size-structured planktonic ecosystem models. *Ecol. Model.* **468**, 109907 (2022).
9. M. Schartau, M. R. Landry, R. A. Armstrong, Density estimation of plankton size spectra: a reanalysis of IronEx II data. *J. Plankton Res*. **32**, 1167–1184 (2010).
10. A. G. Taylor et al., Sharp gradients in phytoplankton community structure across a frontal zone in the California Current Ecosystem. *J. Plankton Res*. **34**, 778–789 (2012).
11. A. Oschlies, P. Brandt, L. Stramma, S. Schmidtko, Drivers and mechanisms of ocean deoxygenation. *Nat. Geosci.* **11**, 467–473 (2018).
12. L. Stramma, S. Schmidtko, L. A. Levin, G. C. Johnson, Ocean oxygen minima expansions and their biological impacts. *Deep-Sea Res. Pt. I.* **57**, 587–595 (2010).
13. R. G. Asch, C. A. Stock, J. L. Sarmiento, Climate change impacts on mismatches between phytoplankton blooms and fish spawning phenology. *Glob. Change Biol.* **25**, 2544–2559 (2019).
14. K. F. Edwards, E. Litchman, C. A. Klausmeier, Functional traits explain phytoplankton community structure and seasonal dynamics in a marine ecosystem. *Ecol. Lett*. **16**, 56–63 (2013).
15. G. M. Hallegraeff, Harmful algal blooms: a global overview. *Manual on harmful marine microalgae*. **33**, 1–22 (2003).
16. P. T. Strub, J. S. Allen, A. Huyer, R. L. Smith, Large-scale structure of the spring transition in the coastal ocean off western North America. *J. Geophys. Res-Oceans.* **92**, no. C2, 1527–1544 (1987).
17. E. Di Lorenzo, et al. Nutrient and salinity decadal variations in the central and eastern North Pacific. *Geophys. Res. Lett.* **36**, no. 14 (2009).
18. P. C. Fiedler, Environmental change in the eastern tropical Pacific Ocean: review of ENSO and decadal variability. *Mar. Ecol. Prog. Ser.* **244**, 265–283 (2002).
19. A. C. Thomas, P. Brickley, R. Weatherbee, Interannual variability in chlorophyll concentrations in the Humboldt and California Current Systems. *Prog. Oceanogr.* **83**, no. 1-4, 386–392 (2009).
20. J. Fiechter, C. A. Edwards, A. M. Moore, Wind, Circulation, and Topographic Effects on Alongshore Phytoplankton Variability in the California Current. *Geophys. Res. Lett.* **45**, 3238–3245 (2018).
21. M. Jacox, C. A. Edwards, Upwelling source depth in the presence of nearshore wind stress curl. *Geophys. Res. Lett-Oceans.* **117**, C5 (2012).
22. N. Gruber et al., Eddy-induced reduction of biological production in eastern boundary upwelling systems. *Nat. Geosci.* **4**, 787–792 (2011).
23. L. Renault et al., Partial decoupling of primary productivity from upwelling in the California Current system. *Nat. Geosci.* **9**, 505–508 (2016).
24. S. J. Lentz, D. C. Chapman, The Importance of Nonlinear Cross-Shelf Momentum Flux during Wind-Driven Coastal Upwelling. *J. Phys. Oceanogr.* **34**, 2444–2457 (2004).
25. X. J. Capet, P. Marchesiello, J. C. McWilliams, Upwelling response to coastal wind profiles. *Geophys. Res. Lett.* **31** (2004), doi:10.1029/2004GL020123.
26. A. Bakun, C. S. Nelson, The Seasonal Cycle of Wind-Stress Curl in Subtropical Eastern Boundary Current Regions. *J. Phys. Oceanogr.* **21**, 1815–1834 (1991).
27. S. A. Henson, A. C. Thomas, Phytoplankton scales of variability in the California Current System: 2. Latitudinal variability. *J. Geophys. Res-Oceans.* **112**, no. C7 (2007).
28. S. A. Henson, A. C. Thomas, Interannual variability in timing of bloom initiation in the California Current System. *J. Geophys. Res-Oceans.* **112**, no. C8 (2007).
29. S. J. Bograd et al., Phenology of coastal upwelling in the California Current. *Geophys. Res. Lett.* **36**, no. 1 (2009).
30. S. A. Henson, I. Robinson, J. T. Allen, J. J. Waniek, Effect of meteorological conditions on interannual variability in timing and magnitude of the spring bloom in the Irminger Basin, North Atlantic. *Deep-Sea Res. Pt. I.* **53, no. 10,** 1601-1615 (2006).
31. M. Jacox, C. Edwards, E. Hazen, S. Bograd, Coastal Upwelling Revisited: Ekman, Bakun, and Improved Upwelling Indices for the U.S. West Coast. *J. Geophys. Res-Oceans.* **123** (2018).
32. P. M. Kosro, W. T. Peterson, B. M. Hickey, R. K. Shearman, S. D. Pierce, Physical versus biological spring transition: 2005. *Geophys. Res. Lett.* **33**, no. 22 (2006).
33. M. Kahru, R. M. Kudela, M. Manzano-Sarabia, B. G. Mitchell, Trends in the surface chlorophyll of the California Current: Merging data from multiple ocean color satellites. *Deep Sea Res. Pt. II.* 77, 89-98 (2012).
34. S. Maritorena, D. A. Siegel, Consistent merging of satellite ocean color data sets using a bio-optical model. *Remote Sens. Environ.* **94, no. 4**, 429-440 (2005).
35. California merged satellite-derived 1-km dataset, (available at http://spg-satdata.ucsd.edu/).
36. M. Kahru, R. Kudela, M. Manzano-Sarabia, B. G. Mitchell, Trends in primary production in the California Current detected with satellite data. *J. Geophys. Res-Oceans.* **114, no. C2** (2009).
37. Access Data | PO.DAAC / JPL / NASA. *Physical Oceanography Distributed Active Archive Center (PO.DAAC),* (available at https://podaac.jpl.nasa.gov/dataaccess).
38. M. García-Reyes, J. L. Largier, Seasonality of coastal upwelling off central and northern California: New insights, including temporal and spatial variability. *J. Geophys. Res-Oceans.* **117, no. C3** (2012).
39. D. A. Siegel, S. C. Doney, J. A. Yoder, The North Atlantic spring phytoplankton bloom and Sverdrup's critical depth hypothesis. *Science.* **296, no. 5568**, 730–733 (2002).
40. S. A. Henson, R. Sanders, C. Holeton, J. T. Allen., Timing of nutrient depletion, diatom dominance and a lower-boundary estimate of export production for Irminger Basin, North Atlantic. *Mar. Ecol. Prog. Ser.* **313**, 73–84 (2006).
41. A. Huyer, Coastal upwelling in the California Current system. *Prog. Oceanogr.* **12, no. 3**, 259–284 (1983).
42. J. Fiechter, C. A. Edwards, A. M. Moore, Wind, circulation, and topographic effects on alongshore phytoplankton variability in the California Current. *Geophys. Res. Lett.* **45**, 3238–3245 (2018).
43. Capet, X. J., P. Marchesiello, J. C. McWilliams, Upwelling response to coastal wind profiles. *Geophys. Res. Lett.* **31**, (2004).
44. M. G. Jacox, C. A. Edwards, Effects of stratification and shelf slope on nutrient supply in coastal upwelling regions. *J. Geophys. Res-Oceans.* **116**, (2011), doi:10.1029/2010JC006547.
45. M. G. Jacox, A. M. Moore, C. A. Edwards, J. Fiechter, Spatially resolved upwelling in the California Current System and its connections to climate variability. *Geophys. Res. Lett.* **41**, 3189–3196 (2014).
46. N. Perlin, E. D. Skyllingstad, R. M. Samelson, Coastal atmospheric circulation around an idealized cape during wind-driven upwelling studied from a coupled ocean–atmosphere model. *Mon. Weather Rev.* **139, no. 3**, 809–829 (2011).
47. F. Kessouri, et al., Coastal eutrophication drives acidification, oxygen loss, and ecosystem change in a major oceanic upwelling system. *P. Natl. Acad. of Sci. USA*. **118, no. 21** (2021).
48. M. Hepner-Medina, SCCOOS | (2021), (available at https://sccoos.org/).
49. CalCOFI-Scripps Institution Of Oceanography, California Current Ecosystem LTER, E. Venrick, Temporal and spatial changes in the abundance and species composition of phytoplankton in the California Current from samples collected aboard CalCOFI cruises between summer 1996 and spring 2012. (2017), , doi:10.6073/PASTA/AD6A9128A40EC0CA443BE4030F115CB4.

Quantifying spatial associations between ^{26}Al and galactic star formation

Sophia J. Westerkamp[1], Lauren Tafla[2], Edward D. Young[2]

[1]Department of Chemistry & Biochemistry. [2]Department of Earth, Planetary, and Space Sciences.

ABSTRACT

The origins of short-lived radionuclides in the early solar system such as ^{26}Al have been debated for decades, and multiple theories have come to the forefront of the discussion. This paper explores evidence for the theory that the abundance of ^{26}Al was typical of star-forming regions similar to those observed in the Galaxy today. As radioactive ^{26}Al decay serves as the main heat source and chemical and mineralogical evolution-shaping mechanism for planetary embryos, modeling the origins of ^{26}Al is essential to understand the early evolution of planets. In this paper, we present a statistical approach that measures and quantitatively describes the association between galactic-scale maps of ^{26}Al and carbon monoxide (CO), a tracer of molecular clouds where star formation exclusively takes place. Using an unweighted area cross-tabulation of CO and ^{26}Al maps, we found that the conditional probability of CO occurring given the presence of ^{26}Al is 0.73, potentially indicating a strong association between the two molecules. Ongoing research seeks to incorporate a correction for the expected degree of matching by using a Monte Carlo simulation of the distributions, as well as addressing the varying degree of pixel intensities in the maps and association between near neighbors. This work has implications for understanding the radiochemistry of the early solar system and evaluating the likelihood of it sharing a similar environment to the self-enriched massive star-forming regions observed in the Galaxy today.

INTRODUCTION

In order to fully understand the formation and evolution of the solar system it is critical to gain a complete picture of the origins of radionuclides such as ^{26}Al, ^{36}Cl and ^{41}Ca. However, the source of these radionuclides is currently unknown and widely debated (1-8). These short-lived radionuclides (SLRs) are formed within the upper hydrogen and helium burning layers of massive stars and within supernovae (9). For ^{26}Al in particular, there exists a discrepancy between its early solar value and its average interstellar medium (ISM) value. Lee et al. 1977 calculated the original concentration of ^{26}Al in the early solar system from known concentrations of the decay product of ^{26}Al, - ^{26}Mg - in meteorites. They found that the solar system formed with an abundance of $n(^{26}Al)/n(^{27}Al) = 5.210^{-5}$ (10). However, Tang et al. 2012 found the average interstellar medium (ISM) value to be ^{26}Al to be $n(^{26}Al)/n(^{27}Al) = 3.010^{-6}$, more than 10 times less than the early solar value calculated by Lee et al (11). Two main theories were introduced to explain this discrepancy. The first posits that the large solar ^{26}Al value is the result of individual progenitor events (1-3), whereas the second theory attributes the large abundance to a heightened concentration of ^{26}Al and other SLRs in star forming regions (SFRs) as opposed to the general ISM (4-8).

According to models of the first theory, close encounters with ^{26}Al sources such as collapsed supernovae (SNe), asymptotic giant branch stars, or Wolf-Rayet (WR) stars caused the heightened proportion of ^{26}Al in early the solar system. Ouellette et al. 2007 and Ouellette et al. 2010 claim that the most likely source for the early solar radionuclide abundances was a nearby supernovae that occurred soon after the formation of the solar system's protoplanetary disk (1,2). Wasserberg et al. 2006 lists irradiation, asymptotic giant branch stars, and SNe as potential sources for various SLR abundances (3). Though different, both of these methods rely upon unique progenitor events occurring during the early development of the protoplanetary disk, implying that the solar system's abundances of SLRs is essentially unique. However, the probability of events such as these occurring such that they mimic the initial solar abundances of SLRs is very low (12). For the early solar system to be adequately enriched with SLRs, Adams et al. estimated that the SN must fall within the range of d ≈ 0.1-0.3 pc (12), however, in order to avoid stripping of the early solar system, the SN must have been around 0.2 pc away from the early solar system (12,13). The range of radii that would allow for both of these conditions to be met is very narrow, and the probability of such an event occurring is only $P_d \approx 0.02$ (12). Other constraints such as the lifetime of the SN progenitor and the radiation produced by a SN further decrease the probability of the early solar system's abundances of SLRs being the result of an individual progenitor event (12).

According to the second theory, star forming regions are heightened in their abundances of SLRs in comparison to the general galactic medium, indicating that the early solar system's abundance of SLRs was usual rather than unique. To test this theory, Jura et al. 2013 studied asteroids accreted by white dwarfs (6). By locating the presence of igneous differentiation in accreted asteroids, they calculated that the value of ^{26}Al necessary to produce this igneous differentiation must have been at least 3×10^{-5} (6). This value is consistent with that of our own solar system (6,10), and leads to the assumption that the enrichment of ^{26}Al in other solar systems could occur if ^{26}Al is concentrated within SFRs (6). The potential source of these heightened abundances is solar winds, particularly those from massive stars such as WR stars. Due to their high masses, WR stars do not travel far from their birth sites before their death. Thus, WR winds containing ^{26}Al may remain largely contained to the SFR where the star originally formed (5). Young 2014 showed that a two-phase ISM model with one dense phase representing SFRs and another representing the general galactic medium accurately describes the early solar system abundances of all considered radionuclides (4). By incorporating WR winds, Young 2014 was able to accurately fit both short- and long-lived radionuclide data to a single model, providing more evidence that the solar abundances of radionuclides could have come from the parental molecular cloud in an SFR (4).

Further supporting the second theory, Young 2016 found strong probabilistic evidence that ^{26}Al exists preferentially in SFRs, suggesting that the solar system formed in a large SFR similar to Cygnus today (5). Finally, Young 2018 analyzed 15 radionuclides and their stable partners. The results suggested that the sun likely formed in a typical SFR and that the apparent enrichment of ^{26}Al in the solar system was due to an under-estimation of the enrichment by WR winds in SFRs (7). All of these sources support the second theory that ^{26}Al lies predominantly in star forming regions, and that the initial radionuclide abundances of the solar system are usual rather than unique. However, while valuable research has been conducted to support the second theory, much of it relies upon small data sets and/or simulations or models to support their arguments. In order to supplement previous research, this paper looks towards the Milky Way Galaxy itself to provide evidence that ^{26}Al resides predominantly within SFRs. Looking at maps of the ^{26}Al distribution across the Galaxy allowed us to provide both quantitative and qualitative evidence that the abundance of ^{26}Al is concentrated within and around SFRs.

If ^{26}Al is shown to reside predominantly within SFRs, the solar system's initial abundances of radionuclides would fall well within the expected values, indicating that it is not unique in its initial radionuclide abundances (6). Understanding whether the solar system formed under ordinary conditions is important in evaluating the possibility of similar planetary systems existing elsewhere in the Galaxy and beyond (7). As ^{26}Al is the main heat source for planetary embryos, modeling the origins of the molecule assists in understanding the early evolution of planets and placing the solar system's origin within context of the rest of the

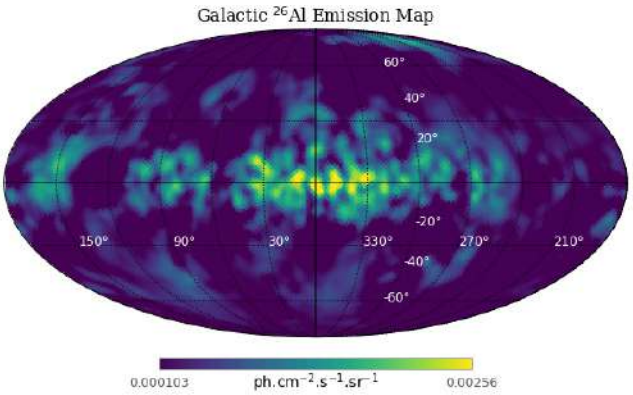


Figure 1. ^{26}Al Galactic Distribution Map. A visualization of the International Gamma-Ray Astrophysics Laboratory (INTEGRAL) 26Al data, gathered by testing for 1.8 MeV gamma ray emission from radioactive 26Al decay across the galactic spectrum. Emission values are given as ph cm-2s-1sr-1. Regions with little to no 26Al present appear as dark blue, whereas regions with high quantities of 26Al appear as yellow. Note that this map has been resized by a factor of 16 to match the pixel count of the CO map seen in Figure 2. A visual concentration of 26Al towards the galactic center and star forming regions is apparent.

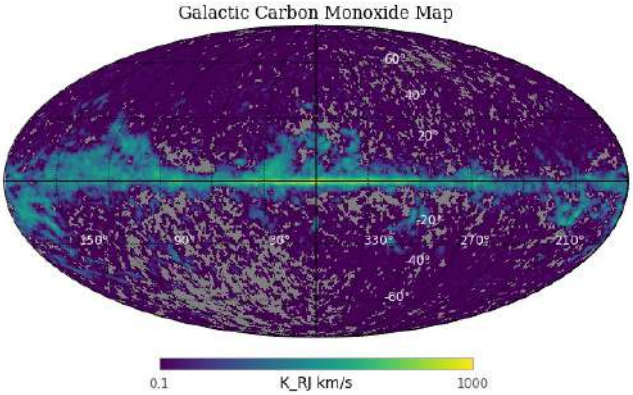


Figure 2. CO Galactic Distribution Map. A visualization of the Planck carbon monoxide (CO) data with rotational transition lines for CO used as CO emission tracers in K km s-1. Regions with little to no CO present appear as dark blue, whereas regions with high quantities of CO appear as yellow. Gray regions represent locations where there is negligible or zero CO. Areas of high CO concentration align with the locations of star forming regions.

Galaxy (7). Finally, the spatial positioning of ^{26}Al relative to SFRs may help to predict how WR stars end their lives: as black holes or as supernovae (4). In summary, identifying the position and source of ^{26}Al within the Galaxy is a necessary step towards understanding the formation of our solar system relative to others as well as understanding the end-life behavior of WR stars.

METHODS

Creating 26Al and CO galactic mollweide projection maps

Data from Planck Collaboration (P. A. R. Ade et al. 2013), which maps the occurrence of CO across the Galaxy and International Gamma-Ray Astrophysics Laboratory (INTEGRAL) data

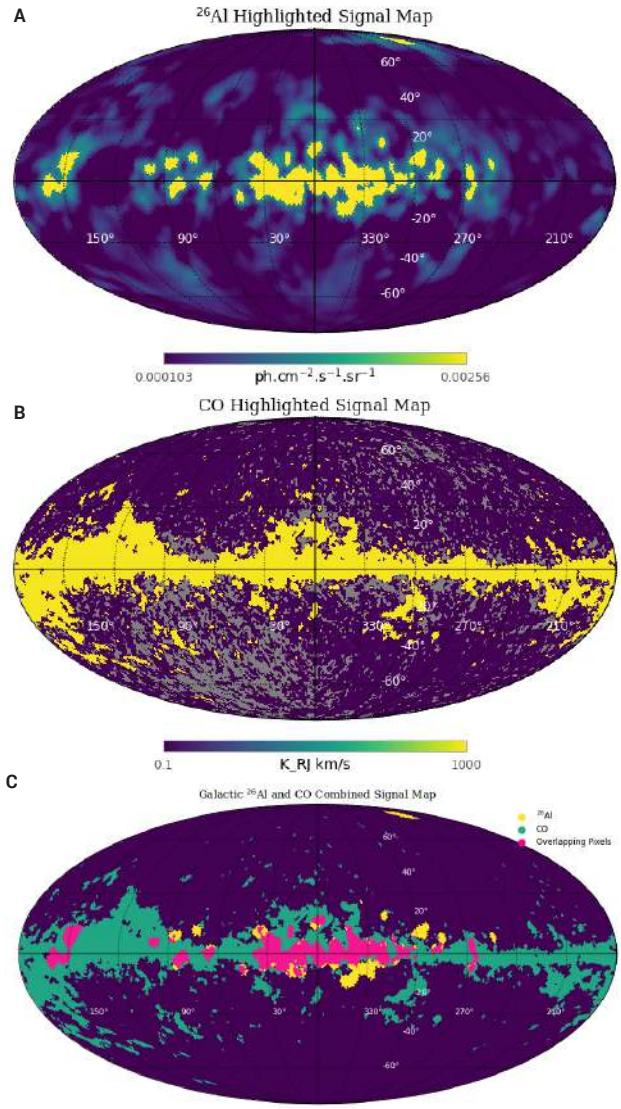


Figure 3A, 3B and 3C. ^{26}Al and CO Highlighted Threshold and Combined Signal Galactic Distribution Maps. A) A visualization of the International Gamma-Ray Astrophysics Laboratory (INTEGRAL) 26Al data, in which all pixels with 26Al values greater than 5.410-4 ph cm-2s-1sr-1 are bright yellow in order to highlight regions containing significant 26Al. **B)** A visualization of the Planck carbon monoxide (CO) data, where all pixels with CO emission greater than 0.2 K km s-1 are bright yellow. All major galactic star forming regions exist in yellow regions. **C)** The accumulation of significant 26Al and CO pixels combined into a single distribution map. Regions absent of significant 26Al or CO levels are indicated in dark purple, and regions containing significant levels of either molecule are indicated in yellow for 26Al and green for CO. Regions with significant levels of both 26Al and CO are indicated in pink.

(Bouchet et al. 2015), which maps the occurrence of ^{26}Al across the Galaxy, were utilized to quantify the correlation between ^{26}Al and star forming regions in the Milky Way (14,15). SFRs occur in areas with high molecular gas concentration, and thus, high density locations of molecules such as carbon monoxide, neutral hydrogen, and ammonia are all considered molecular tracers for SFRs. Carbon monoxide was chosen to study due to its high-quality and readily-available galactic abundance and positionings data. A large correlation between the spatial positionings of ^{26}Al and CO across the Galaxy would provide evidence that ^{26}Al predominantly resides in SFRs rather than uniformly throughout the ISM.

Galactic positionings and abundances of ^{26}Al were determined through measurement of the gamma-ray 1.809 MeV line emission associated with ^{26}Al decay. CO data was gathered by the Planck Collaboration (14) through measurement of the CO rotational line emission. Both sources of data were contained within Flexible Image Transport System (FITS) files. To convert the data into images, code was created with the Python programming language using the package astropy.io.fits to convert the tabulated data into three dimensional graphical representations (16,17). These representations can be seen in Figures 1 and 2, which depict the galactic abundances and positionings of ^{26}Al in units of ph cm-2s-1sr-1; and the galactic abundances and positionings of CO in units of K km s-1, respectively. In these two figures, a visual correlation appears between ^{26}Al and CO across the galactic center as areas of high 26Al density seem to visually align with areas of high CO density. If 26Al was consistent across the ISM, there would be no concentration of 26Al towards regions of high CO density beyond the general galactic plane.

Creating an ^{26}Al and CO combined signal map

Originally, the CO map contained 786,432 pixels whereas the ^{26}Al map contained only 49,152 pixels, a difference due to measurement techniques used to gather each data set. As ^{26}Al was measured via its decay emission energy and CO was measured through its rotational emission, both data sets contain a separate quantity of pixels with separate units. In order to run a direct mathematical comparison of the two maps, it was first necessary to resize the ^{26}Al map to fit the pixel count of the CO map. This was completed in Python by reshaping the original data array to contain 786,432 pixels rather than 49,152 pixels. In essence, this code spatially decomposed each ^{26}Al pixel into 16 smaller individual pixels.

To determine which ^{26}Al map pixels contained a significant amount of ^{26}Al, a threshold value of 5.410^{-4} ph cm^{-2} s^{-1} sr^{-1} was set. This was chosen as it is the same vale used by Bouchet et al. 2015, and highlights all regions that visually appear to contain significant amounts of ^{26}Al (15). A new ^{26}Al map was subsequently created within Python which set all values above this threshold to a standard color (yellow), shown in Figure 3. To determine which CO pixels contained a significant amount of CO, a threshold of value of 0.2 K km s^{-1} was set such that all known galactic SFRs, particularly the Orion region, were included above the CO threshold. To determine this value, we visually compared the CO map with locations of known SFRs throughout the Galaxy. All pixels which contained CO data above the threshold value were then set to a standard color (yellow) as can be seen in Figure 3.

In order to quantitatively compare the positions of CO and ^{26}Al in the Galaxy, a final association map was created where only ^{26}Al and CO data that were above the individual thresholds were considered. Individual arrays were created within Python for four cases: ^{26}Al pixels below the threshold, ^{26}Al pixels above the

threshold, CO pixels above the threshold, and pixels for which both ^{26}Al and CO were above the threshold. The entry values for each array were then set to different numbers corresponding to specific points on the color scale. By doing so, the pixels in each array were set to a specific color within the map. The final map showcasing the positions of all overlapping ^{26}Al and CO pixels above their respective threshold values can be seen in Figure 3.

Conditional probability calculations

To determine the statistical significance of ^{26}Al and CO overlap, a conditional probability calculation was performed by calculating the probability of CO occurring given the presence of ^{26}Al, written as P(CO | ^{26}Al). From probability theory, this conditional probability is defined as

$$P(CO \mid {}^{26}Al) = \frac{P(CO \cap {}^{26}Al)}{P({}^{26}Al)} \quad \text{(Eq.1)}$$

where P(CO ⋂^{26}Al) is the probability of both CO and ^{26}Al occurring together, and P(^{26}Al) is the probability of just ^{26}Al occurring. These values were determined from pixel counts of CO, ^{26}Al, and concurrence of CO and ^{26}Al.

RESULTS

Unweighted spatial association analysis using binary maps

Visually, a qualitative correlation can be seen between the ^{26}Al and CO maps as illustrated in Figures 1, 2, and 3. To quantify this overlap, an unweighted spatial association analysis using the ^{26}Al and CO maps was used. This simple analysis began by counting the number of CO and ^{26}Al pixels that directly overlap in Figure 3, and the percent overlap was calculated through Eq. 2:

$$\frac{\text{Number of pixels of } {}^{26}\text{Al and Co that directly overlap}}{\text{Total number of Al pixels}} \times 100\% \quad \text{(Eq.2)}$$

= Percent of ^{26}Al pixels that overlap with CO

The percent overlap was found to be 73%.

To further analyze the spatial associations between CO and ^{26}Al, a cross-tabulation pixel count was performed to create a two-dimensional summary of the overlap of all possible combinations of the ^{26}Al and CO input maps. To gather this data, the number of pixels that contained both CO and ^{26}Al, either CO or ^{26}Al, or neither molecule were counted, which provided the numerical data which was used in each unconditional and conditional probability calculation described in this paper. Via application of this table, it was calculated that there is three times as much ^{26}Al inside regions containing CO than in regions without CO.

Conditional probability calculations

A central goal of this study was to examine how well the presence of CO can be predicted from the presence of ^{26}Al. Thus the conditional probability of CO given ^{26}Al is a key result. As depicted in Table 1, the unconditional probability of ^{26}Al occurrence is 0.0474, and the unconditional probability of CO occurring at any location on the map, P(CO), was calculated to be ~0.2. The conditional probability of CO occurring given the presence of ^{26}Al is much higher than both of these unconditional probabilities, about 0.7299. . Since the unconditional probability of CO is significantly less than 1, it can be concluded that overlap between ^{26}Al and CO is statistically significant, rather than solely the result of a large P(CO).

Conditional Probability Calculations

P(26-Al)	0.0474
P(CO|26-Al)	0.7299

Table 1. Conditional Probability Calculations Summarizing Table. A summary of the important probabilities calculated from direct pixel count data. The first row depicts the independent probability of 26Al occurrence. The second row depicts the conditional probability of CO occurrence given the presence of 26Al. These probabilities are not depicted as percentages. Calculation of the second probability was completed via the application of equation 1.

DISCUSSION

Although promising research has worked towards demonstrating that ^{26}Al originates within SFRs, there is a lack of quantitative evidence to support these claims. Visually, there is a qualitative correlation between the ^{26}Al and CO maps as seen in Figures 1, 2, and 3. The percent overlap shown in the correlation map marking the positions of ^{26}Al and CO galactic overlap was found to be 73%. This is a promising indicator that much of the galactic supply of ^{26}Al resides within SFRs and supports the theory that the original reserve of ^{26}Al in our solar system is commonplace. Additionally, the finding that ^{26}Al occurs roughly 3 times as often within SFRs as compared to outside of SFRs indicates that the majority of ^{26}Al in the Milky Way is present within SFRs. This provides further evidence towards a common origin theory for the Solar System. Similarly, the sharp increase from 0.0474 to 0.7299 between the unconditional probability of ^{26}Al occurrence and the conditional probability of CO occurrence given the presence of ^{26}Al indicates a strong association between the two. However, we do acknowledge that the expected concentration of both molecules along the galactic plane could be increasing the calculated overlap between CO and ^{26}Al. The work presented in this paper indicates that ^{26}Al is likely concentrated towards SFRs, providing evidence towards the second theory that our solar system is commonplace in its initial abundances of SLRs rather than unique.

In future research, several measures will be taken to improve the initial findings of this work. One area where it may be improved is via the consideration of ^{26}Al that does not directly overlap, but is still close to CO regions as ^{26}Al has the possibility of spreading outside of its parent molecular cloud. This would cause ^{26}Al to be present slightly beyond locations of CO. Thus, when determining the spatial correlation between ^{26}Al and CO,

1	4	7	4	1
4	16	26	16	4
7	26	41	26	7
4	16	26	16	4
1	4	7	4	1

Table 2. Conditional Probability Calculations Summarizing Table. A summary of the important probabilities calculated from direct pixel count data. The first row depicts the independent probability of 26Al occurrence. The second row depicts the conditional probability of CO occurrence given the presence of 26Al. These probabilities are not depicted as percentages. Calculation of the second probability was completed via the application of equation 1.

it would be ideal to obtain a weighted correlation between the two maps in which ^{26}Al that is nearby, but not fully overlapping CO regions is considered. This will be accomplished by applying a weighted statistical approach to both maps that takes into account near neighbors along with varying pixel intensity. A visual example of a weighted gaussian distribution of pixels that may be used is shown in Figure 4. A Monte Carlo (MC) simulation will be performed with an autocorrelation correction and an established local domain to create a cross-tabulation that effectively averages 1000 simulated maps. This MC simulation will output expected values for the locations of ^{26}Al and CO to be used for comparison. In order to take into account spatial autocorrelation, the numbers of ^{26}Al and CO will be kept constant in the MC simulations, essentially allowing only the locations to vary. In general, with a greater number of samples introduced into a spatial dataset and the closer together they are, the more they will be spatially autocorrelated. In order to establish a local domain to make measurements in, the variation of "contrast" with distance will be calculated as described in Bonham-Carter 1994 (18). Comparing these values with the observed values determined through pixel counts will result in a weighted statistical analysis of the spatial positioning of ^{26}Al relative to CO. This approach is currently being investigated and will be addressed in a future work.

If ^{26}Al resides predominantly within SFRs, it is likely that other short-lived radionuclides such as ^{36}Cl and ^{41}Ca may be preferentially concentrated towards SFRs as well, implying that the observed high abundances of these SLRs in the solar system may also be explained by the second theory. However, more research is needed to quantify this possibility. The implication that the solar system is normal in its initial radionuclide abundances suggests that similar solar systems may exist in SFRs elsewhere in the Milky Way Galaxy. Similarly, if our solar system's abundance of ^{26}Al is commonplace, it is likely that our solar system was formed via typical rather than abnormal means, helping to place the solar system's origins within the context of the rest of the galaxy. Studying the spatial positioning of ^{26}Al relative to SFRs is a valuable step towards understanding the formation history of our solar system and how similar solar systems may exist elsewhere in the galaxy and beyond.

ACKNOWLEDGMENTS

The author would like to thank Lauren Tafla for her hard work towards the research presented in this paper as well as her thoughtful and supportive input throughout the writing process. The author would also like to thank Professor Edward D. Young for his input regarding the future of this research and his continued support throughout the project.

REFERENCES

1. N. Ouellette, S. J. Desch, J. J. Hester, Interaction of supernova ejecta with nearby protoplanetary disks. *ApJ.* **662,** 1268-1281 (2007).
2. N. Ouellette, S. J. Desch, J. J. Hester, Injection of supernova dust in nearby protoplanetary disks. *ApJ.* **711,** 597-612 (2010).
3. G. J. Wasserburg, M. Busso, R. Gallino, K. M. Nollett, Short-lived nuclei in the early solar system: possible AGB sources. *Nucl. Phys. A.* **777,** 5-69 (2006).
4. E. D. Young, Inheritance of solar short- and long-lived radionuclides from molecular clouds and the unexceptional nature of the solar system. *EPSL.* **392,** 16-27 (2014).
5. E. D. Young, Bayes' theorem and early solar short-lived radionuclides: the case for an unexceptional origin for the solar system. *ApJ.* **826,** 129 (2016).
6. M. A. Jura, S. Xu, E. D. Young, 26Al in the early solar system: not so unusual after all. *ApJ.* **775,** L41 (2013).
7. E. D. Young, The birth environment of the solar system constrained by the relative abundances of the solar radionuclides. *IAU.* **14**, 70-77 (2018).
8. Y. Fujimoto, M. R. Krumholz, S. Tachibana, Short-lived radioisotopes in meteorites from Galactic-scale correlated star formation. *Mon. Notices Royal Astron. Soc.* **480,** 4025-4039 (2018).
9. B. S. Meyer, Synthesis of short-lived radioactivities in a massive star. *ASPCS.* **341**, 515-526 (2005).
10. T. Lee, D. A. Papanastassiou, G. J. Wasserburg, Aluminum-26 in the early solar system: fossil or fuel? *ApJ.* **211**, L107-L110 (1977).
11. H. Tang, N. Dauphas, Abundance, distribution, and origin of 60Fe in the solar protoplanetary disk. *EPSL.* **359-360,** 248-263 (2012).
12. F. C. Adams, The birth environment of the solar system. *Annu. Rev. Astron. Astrophys.* **48,** 47-85 (2010).
13. R. A. Chevalier, Young circumstellar disks near evolved massive stars and supernovae. *ApJ.* **538,** L151-L154 (2000).
14. Planck Collaboration: P. A. R. Ade, et al., Planck 2013 results. XIII. Galactic CO emission. *Astron. Astrophys.* **571,** A13 (2014).
15. L. Bouchet, E. Jourdain, J. Roques, The galactic 26Al emission map as revealed by integral SPI. *ApJ.* **801,** 142 (2015).
16. T. Robitaille, et al., Astropy: A community Python package for astronomy. *Astron. Astrophys.* **558,** A33 (2013).
17. Astropy Collaboration: A. Price-Whelan, et al., The Astropy Project: Building an inclusive, open-science project and status of the v2.0 core package. *Cornell University.* **1,** 1-39 (2018).
18. G. F. Bonham-Carter, Geographic Information Systems for Geoscientists: Modeling with GIS. Pergamon, Oxford. (1994).

Assessing enzalutamide-induced metabolic rewiring in prostate cancer

Jazmin Michel Mondragon[1], *Andrew Goldstein*[1] ,*Matthew Bernard*[1]

[1]Department of Molecular, Cell, and Developmental Biology.

ABSTRACT

Prostate cancer is the most commonly diagnosed cancer among men and one of the leading causes of cancer-related death. Localized prostate cancer treatment involves surgical removal of the prostate, a highly successful procedure, but recurrent prostate tumors could lead to patient lethality. Current treatment of recurrent prostate cancer involves targeting the androgen receptor signaling axis in cancerous cells. However, some tumors are capable of developing resistance to these therapies. Long-term treatment with the clinical anti-androgen drug enzalutamide often leads to treatment resistance. As tumor cells adapt to anti-androgen therapy, they rewire their metabolism, suggesting that metabolic pathways may be critical for treatment-resistance and are thus potential druggable targets. Here, alterations in cellular metabolism that accompany this gain of treatment-resistance were assessed. Expression of α-ketoglutarate, a tricarboxylic acid (TCA) cycle metabolite, has previously been identified to be significantly altered in anti-androgen resistant prostate cancer cells. The goal of this study was to investigate how prolonged enzalutamide treatment alters enzymes that metabolize α-ketoglutarate to assess their role in treatment resistance. Using metabolic profiling and western blotting, the role of key metabolic pathways, specifically those involved with α-ketoglutarate, in prostate cancer treatment-resistance were investigated. The functional role of key metabolic enzymes were also evaluated through genetic and pharmacologic analysis. Preliminary results showed that oxoglutarate dehydrogenase-like, a metabolic enzyme involved in the TCA cycle, had significant changes in protein and mRNA expression in response to enzalutamide treatment. Identifying targetable metabolic pathways allows for a deeper understanding of metabolism in cancer cells and helps to identify novel treatment strategies for treatment-resistant prostate cancer.

INTRODUCTION

Prostate cancer is a common cause of lethality in male-born patients. Localized prostate cancer therapeutics involve active surveillance, local radiotherapy, or surgical removal of the prostate (*1*). However, prostate cancer often spreads and becomes recurrent, making it more difficult to treat. Recurrent prostate cancer is currently treated by targeting the androgen receptor (AR) signaling axis with androgen deprivation therapy (ADT), which lowers circulating androgen levels (*1*). These recurrent tumors have been found to develop resistance to androgen-targeting therapies and progress to more aggressive prostate cancer (*1*). Thus, new strategies to target the androgen signaling axis are needed or new methods to re-sensitize prostate cancer tumors to androgen therapies for effective treatment.

Enzalutamide is a clinical second generation, anti-androgen drug that can be used to treat ADT-resistant prostate tumors (*1,2*). These anti-androgens inhibit the binding of androgens to their corresponding ARs, prohibiting the translocation of AR into the nucleus and resulting in the inability of ARs to activate androgen target genes that promote tumor growth, which eventually causes tumor death (*1,2*). Treatment with enzalutamide often leads to resistant cancer cell phenotypes including that of neuroendocrine prostate cancer (NEPC) (*2*). These tumor cells have the ability to adapt to ADT by rewiring their metabolism (*2*).

Tumor cells rewire their metabolism to meet bioenergetic demands and evade therapeutic pressures (*1*). This indicates that metabolic pathways may be crucial for treatment-resistance and are potential druggable targets. Therein, changes in metabolic enzyme expression in models of treatment-tolerant prostate cancer cells using RNA sequencing were evaluated *in vitro*. Notably, there was a marked change in the transcriptional expression of oxoglutarate dehydrogenase-like (OGDHL), a tricarboxylic acid (TCA) cycle enzyme. The TCA cycle generates NADH and FADH2 in order to facilitate oxidative metabolism and fuel cellular bioenergetic demands. OGDHL is a neuronal-specific variant of oxoglutarate dehydrogenase (OGDH), which plays a role in the α-ketoglutarate dehydrogenase complex (*3*). This complex catalyzes the reaction of converting α-ketoglutarate to succinyl-CoA within the TCA cycle (*3*). Previous studies have shown altered *OGDHL* expression in certain cancers, such as liver cancer, and have indicated it may be a prognostic biomarker (*4*). In hepatocellular carcinoma, silencing of *OGDHL* demonstrated an upregulation of the α-ketoglutarate to citrate ratio and a decrease in survival rate (*5*). Thus, determining the relationship be-

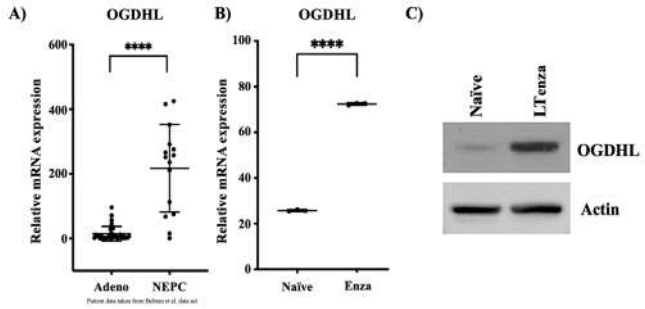


Figure 1. Cellular phenotype and enzalutamide treatment affect Oxoglutarate Dehydrogenase-Like (OGDHL) levels. (A) Previous results identified increased differential mRNA expression of OGDHL at the transcriptional level in neuroendocrine prostate cancer (NEPC). The RNA-sequencing patient data was taken from Beltran et al. data set (6). **(B)** Increased transcriptional levels of OGDHL expression were seen after enzalutamide treatment in patient samples of castrate-resistant prostate cancer (CRPC). **(C)** Increased protein expression of OGDHL was seen in NEPC 16D cell line long-term enzalutamide cells compared to NEPC 16D treatment-naive cells. Two-tailed t-test with Welch's correction. **** indicate p<0.0001. Wnt target genes is initiated (10).

tween treatment-resistant prostate cancer, α-ketoglutarate, and OGDHL is important for a deeper understanding of the role of OGDHL in cancer cell metabolic rewiring and re-sensitization to anti-androgen therapies.

OGDHL exhibits a robust increase in transcriptional and translational expression across models of NEPC (Figure 1) (*6*). Relative mRNA expression of *OGDHL* increased in clinical NEPC patients and after enzalutamide treatment *in vitro* (Figure 1A, B) (*6*). Similarly, following long-term enzalutamide treatment, there is high protein expression of OGDHL, it was hypothesized that OGDHL has a potential role in the development of enzalutamide resistance in recurrent prostate cancer (Figure 1C) (6). Here, the kinetics of *OGDHL* expression elevation after exposure to enzalutamide treatment was investigated and cell line models (LNCaP and 16D) to assess the role of key enzymes (OGDH and OGDHL) in enzalutamide resistance using genetic-based approaches were created.

MATERIALS AND METHODS

Cell lines

Two human prostate cancer cell lines were used: LNCaP, derived from human metastatic prostate cancer, and 16D, derived from a LNCaP xenograft resistant to castration. Cells were grown in RPMI 1640 Medium (Gibco:22400-105) supplemented with 10% fetal bovine serum (Sigma: F0926) and 1% Penicillin-Streptomycin (Gibco:15-140-122).

8-week enzalutamide treatment

16D cells were grown at 37°C in 5% CO2 and passaged when 80-90% confluent. Cells were grown in 10 cm Fisher brand surface treated tissue culture dishes (Fisher: FB012924) in RPMI 1640 Medium (Gibco:22400-105) supplemented with 10% Fetal Bovine Serum (Sigma: F0926 and 1% Penicillin-Streptomycin (Gibco:15-140-122). Cell media contained dimethyl sulfoxide (DMSO) for the control group and enzalutamide (Selleck Chemicals: S1250) dissolved in DMSO (Fisher: BP231-100) for the treatment group at a final concentration of 10 μM. Cells were imaged to observe cell growth in both groups. Cells from each group were harvested approximately each week for further analysis and experiments.

Imaging and Analysis

During long-term enzalutamide treatment, live cells were imaged weekly. Dishes were placed under an inverted microscope (Zeiss Axio Vert.A1) and imaged. Images were taken using a Zeiss Axiocam 506 camera.

Knockdown of OGDH and OGDHL

Lentiviral vectors harboring short hairpin RNA (shRNA) constructs against OGDH and OGDHL were procured from VectorBuilder. *OGDH* in both LNCaP and 16D cells was knocked down. To create the knockdowns, shOGDH vector was added to a 10 cm well plate of LNCaP cells and shOGDHL vector was likewise added to a 10 cm well plate of long-term enzalutamide treated 16D cells. For the control group, lentiviral transduction of shScrambled control was added to LNCaP cells. Viral vectors for both knockdowns were constructed by VectorBuilder. Concentrated lentiviral vector containing shRNA was used to infect cells. Western blot analysis was performed to verify both knockdowns and to assess lineage markers. Knockdowns were also used to conduct metabolic profiling and to determine changes in enzalutamide response.

Western blot analysis

Lysates were harvested in Pierce RIPA Buffer (Thermo Scientific: PI89901), sonicated, and then allowed to rest on ice for 45 minutes before being stored at -20°C. For western blot analysis, the samples loaded consisted of protein lysate, NuPAGE LDS sample buffer (NP0007) and NuPAGE Sample Reducing Agent (NP0004) to a final concentration of 1X. Then heated to 70°C for 15 minutes to denature. Gels used for western blot were NuPAGE 4-12% Bis-Tris, 1.0mm 12-well (Invitrogen: NP0322BOX). The running buffer solution consisted of NuPAGE MOPS SDS Running Buffer (NP0001) at a 1X concentration and was supplemented with NuPAGE Antioxidant (Invitrogen: NP0005). Western blots were run at 200V for 55 minutes. For the transfer of the protein to the membrane, Immobilon-P PVDF transfer membranes (Millipore-Sibma IPVH00010) and 1X NuPAGE Transfer Buffer (NP00061) in 10% Methanol were used. The proteins were transferred to PVDF at 30V for 60 minutes. Membranes were then blocked for 1 hour at room temperature in 5% milk and 1X PBS with 0.1% Tween-20 (Fisher: BP337-500) and incubated with primary antibodies overnight at 4°C. The antibodies used were actin (Invitrogen: 15G5A11/E2), with a dilution of 1:1500; NSE (Proteintech: 66150-1-1g), with a dilution of 1:5000; OGDHL (Invitrogen: PA5-62626); OGDH (Proteintech: 15212-1-AP); and PSA (Cell Signaling 5877S), with dilution factor of 1:1000.

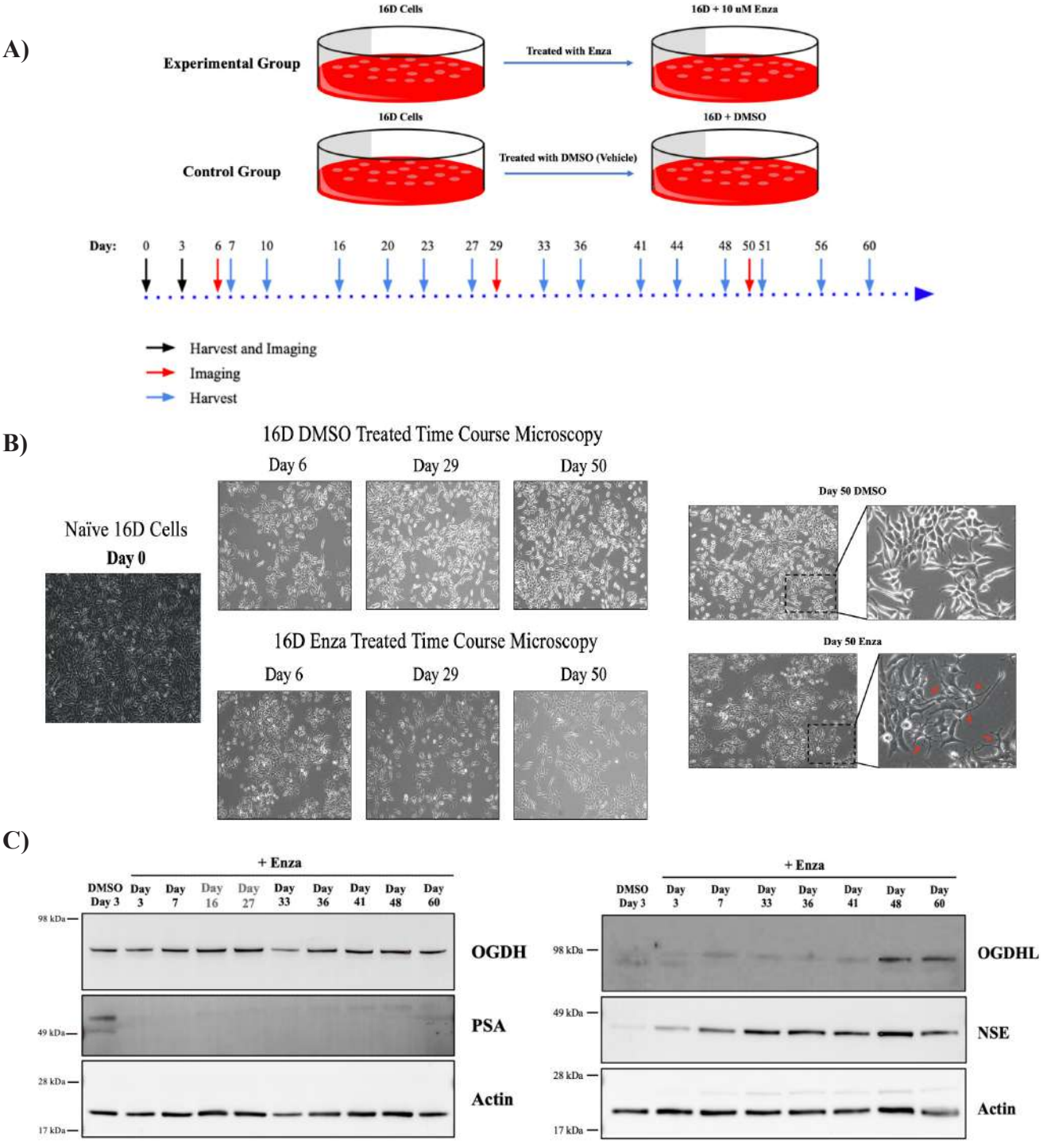


Figure 2. In prolonged enzalutamide treatment, 16D cell lines develop a neuroendocrine phenotype and show increased Oxoglutarate Dehydrogenase-Like (OGDHL) expression. (A) Schematic showing 8-week-long experiment treating 16D cells with enzalutamide (10 μM). The control group was treated with dimethyl sulfoxide (DMSO). Red arrows indicate when cells were imaged and blue arrows indicate when cells that were harvested for lysates. Black arrows indicate cells that were both imaged and harvested. **(B)** 16D cells treated with long-term enzalutamide develop neurite-like projections resembling a more neuroendocrine phenotype, compared to DMSO treated cells. Red arrows in the rightmost panels show neurite-like projections. **(C)** Western blots of DMSO and enzalutamide-treated 16D cell lysates show no significant change in the protein expression level of OGDH in enzalutamide-tolerant cells. Decreased protein levels of prostate-specific antigen (PSA), a target of androgen signaling, verify the function of enzalutamide since PSA is expected to decrease in presence of enzalutamide. Actin was used as a loading control. (D) Western blot DMSO and enzalutamide treated 16D cell lysates shows increased expression of OGDHL and neuron-specific enolase (NSE) associated with prolonged enzalutamide treatment.

RESULTS

16D cells developed a neuroendocrine phenotype in prolonged enzalutamide treatment and showed increased expression of the (OGDHL)

A time-course experiment and western blot analysis were used to analyze changes in androgen-responsive prostate cancer 16D cells treated with prolonged enzalutamide over time (Figure 2A). Images of 16D cells through the 8-week time course experiment show differences between the DMSO treated vehicle control group and the enzalutamide treated experimental group (Figure 2B). Naïve cells at Day 0 were plated in conditions containing either enzalutamide or DMSO. Cells treated with enzalutamide developed neurite-like projections reminiscent of neuroendocrine prostate cancer (NEPC) cells which were most prominent in the last week of enzalutamide treatment (Figure 2B). Western blot analysis was conducted with lysates harvested each week throughout the experiment. Prostate-specific antigen (PSA), a target of androgen signaling which decreases in concentration when enzalutamide is present, was followed through this analysis to ensure enzalutamide was functioning properly, since PSA decreases when enzalutamide is present (Figure 2C). PSA expression levels decreased starting week 1 of enzalutamide treatment and remained low throughout the time course of the experiment (Figure 2C). Protein expression of OGDH remained unchanged with prolonged enzalutamide treatment (Figure 2C). However, with long-term enzalutamide treatment, the expression of *OGDHL* increased (Figure 2D). The levels of OGDHL start low but increase at day 48 and remain high throughout the end of the experiment (Figure 2D). The expression of neuron-specific enolase (NSE), a neural associated gene, also increased with long-term enzalutamide treatment (Figure 2D).

OGDH and OGDHL knockdowns affect the expression of NSE

A genetics-based knockdown of *OGDH* in LNCaP cells and *OGDHL* in 16D cells was conducted (Figure 3A). *OGDHL* expression only increased in response to long-term enzalutamide treatment. Decreased expression of *OGDH* in LNCaP cells is reflected in western blot analysis for shOGDH knockdown groups; additionally, NSE levels decreased slightly (Figure 3B). Western blot analysis also confirms *OGDHL* knockdown in the shOGDHL group for 16D cells (Figure 3C). In addition to the decrease of expression of *OGDHL*, the western blot also shows a greater decrease in NSE compared to the *OGDH* knockdown, thus suggesting that there is an association between the loss of OGDHL and NSE (Figure 3C). Decrease of NSE expression with loss of OGDHL could indicate a loss of the neuroendocrine phenotype seen in 16D cells with prolonged enzalutamide treatment (Figure 2, 3C).

DISCUSSION

Recurrent prostate cancer cells develop resistance to prolonged enzalutamide treatment. Long-term treatment of the anti-androgen drug, enzalutamide, was observed to change the expression of certain metabolites in prostate cancer cell lines. Previous mRNA and protein analysis of the metabolic enzyme OGDHL revealed that long-term treatment with enzalutamide increases the expression of *OGDHL*. Similarly, western blot analysis revealed that OGDHL is the metabolite that demonstrates a consistent pattern of change in prostate cancer cells when exposed to enzalutamide treatment. This work expands upon knowledge on the role of OGDHL in metabolic rewiring and exposes new strategies for determining enzalutamide treatment resistance.

NSE expression increase with prolonged enzalutamide treatment indicates a shift towards a neuroendocrine phenotype that is consistent with the morphological changes seen in the 16D cells. Thus, these results suggest that the OGDHL metabolic enzyme is a key factor that is associated with phenotypic changes of prostate cancer to NEPC cancer after prolonged enzalutamide treatment.

Changes in transcriptional and protein level expression of the metabolic enzyme OGDHL when treated with enzalutamide were demonstrated. Additionally, genetic models to assess how alterations in *OGDH* and *OGDHL* expression influence cell metabolism and response to anti-androgen treatment were generated. *OGDH* was knocked down in LNCaP cells, which are epithelial cells derived from a human prostate carcinoma that use the androgen hormones to upregulate gene targets; this means that LNCaP cells are androgen-sensitive. The *OGDHL* knockdown was conducted in 16D cells, which are derived from the LNCaP cell line; however, they are not androgen-sensitive but are androgen-responsive, which means they are not reliant on androgen hormones to maintain cell function and growth. Androgen-responsive cells do not need androgens to be present to proliferate but still rely on AR signaling. These 16D cells treated long-term in enzalutamide are enzalutamide resistant. *OGDHL* was not knocked down in LNCaP cells because *OGDHL* is not expressed in treatment naïve LNCaP cells (LNCaP); and *OGDH* was not knocked down in long-term enzalutamide 16D cells because the knockdown in the naïve cells was not consistent.

The results show that treating prostate cancer cells with enzalutamide over an 8-week period induces a neuroendocrine phenotype and an increased expression of *OGDHL* in the cells (Figure 2B). However, despite OGDHL being functionally redundant with OGDH, a compensatory downregulation of *OGDH* expression in these treatment-tolerant cells was not observed (Figure 2C). This suggests that OGDHL may have additional functions in treatment-resistant prostate cancer cells. Furthermore, *OGDH* and *OGDHL* were genetically knocked down to assess changes in cell metabolism and disease progression. Through this, NSE, a marker associated with NEPC, appears to decrease slightly with the knockdown of *OGDH*. A greater decrease in NSE was seen in the knockdown of *OGDHL* suggesting a possible loss of the enzalutamide-induced neuroendocrine phenotype in the cells (Figure 3B, C). The greater decrease of NSE when *OGDHL* was knocked down suggests an association between NSE and OGDHL. Together, these results suggest that OGDHL could be a critical enzyme in the rewiring of metabolism occurring in treatment-resistant prostate cancer cells. Other studies have iden-

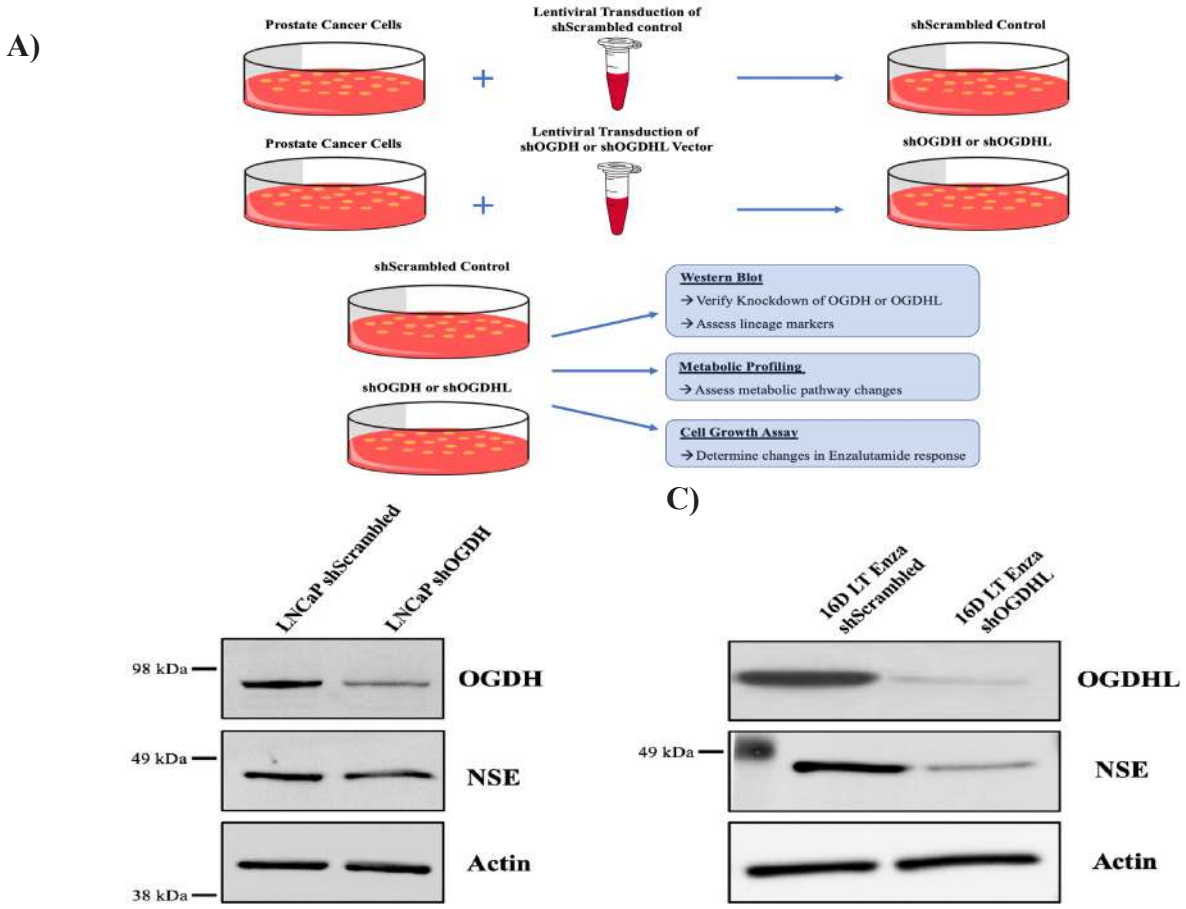


Figure 3. Oxoglutarate Dehydrogenase (OGDH) and OGDHL can be knocked down with a lentiviral vector. A) Schematic showing the genetics-based knockdown of OGDH and OGDHL. **(B)** Western blot verifies protein level knockdown of OGDH in the LNCaP shOGDH (knockdown of OGDH) cell line. Expression of neuron-specific enolase (NSE) slightly decreases in shOGDH. **(C)** Western blot analysis shows decreased protein expression of OGDHL in the 16D long-term enzalutamide (LT Enza) shOGDHL group confirming knockdown of the enzyme. NSE protein expression levels decrease in 16D long-term (LT) shOGDHL. Actin was used as a loading control for both Gels B and C.

tified a similar expression pattern of OGDHL with prolonged enzalutamide treatment. Though transcriptional silencing of *OGDHL* has been previously associated with decreased patient survival in hepatocellular carcinoma, these findings suggest that activation of *OGDHL* may be associated with disease progression in treatment-resistant prostate cancer (5). However, the functional role of OGHDL in prostate cancer progression remains unclear. Therefore, future studies are required to characterize the kinetics of OGDHL induction at the transcriptional and protein levels, assess potential epigenetic regulation of metabolic enzymes in response to prolonged enzalutamide treatment, and identify changes in other metabolic enzymes contributing to changes in cellular metabolism in enzalutamide-resistant cancer cells. These future studies will lead closer to identifying novel therapeutic treatments for anti-androgen drug resistant prostate cancer.

ACKNOWLEDGMENTS

The author thanks everyone in the Goldstein Lab for their support. This work is funded in part by the National Institutes of Health and National Institute of General Medical Sciences through a training grant (NIH MARC T34 GM008563, PI: M. M. McEvoy).

REFERENCES

1. K. Mamouni, G. Kallifatidis, B.L. Lokeshwar, Targeting mitochondrial metabolism in prostate cancer with triterpenoids. *Int J Mol Sci.* **22(5)**, 2466 (2021).
2. S. Bhattacharya, M. Hirmand, D. Phung, S. van Os, Development of enzalutamide for metastatic castration-resistant prostate cancer. *Ann NY Acad Sci.* **1358**, 13-27 (2015).
3. M. Pithukpakorn, Disorders of pyruvate metabolism and the tricarboxylic acid cycle. *Mol Genet Metab.* **85(4)**, 243-6 (2005).
4. Y. Jiao et al., OGDHL expression as a prognostic biomarker for liver cancer patients. *Dis Markers.* **9037131** (2019).
5. W. Dai et al., OGDHL silencing promotes hepatocellular carcinoma by reprogramming glutamine metabolism. *J Hepatol.* **72(5)**, 909-23 (2020).
6. H. Beltran et al., Divergent clonal evolution of castration-resistant neuroendocrine prostate cancer. *Nat Med.* **22(3)**, 298-305 (2016).

Analytical solution of the 1D TDSE for strong laser field electron emission

Benjamin Wang[1], Joshua Mann[1], James B. Rosenzweig[1]

[1]Department of Physics and Astronomy

ABSTRACT

This paper investigates the possibility of using a recently derived analytical solution to the 1-D time-dependent Schrödinger equation to efficiently generate yield curves for laser-induced strong field electron emission from a metallic surface. Strong laser field emission from metallic tips is important because it can create bright, coherent electron beams that may be used for various purposes such as electron microscopy. The analytical solution of the time-dependent Schrödinger equation is numerically evaluated and the results are compared to those obtained from the direct application of numerical methods such as the split-step (Fourier) and Crank-Nicolson methods to similar systems. Yield curves are plotted using both the analytical and numerical results. It is discussed how these yield curves exhibit the theoretically predicted yield scaling (power law with electric field amplitude) as well as channel closing, and these observed characteristics are confirmed against other results in the literature. The difficulties encountered stemming from the high computational complexity of using the analytical solution for yield scaling calculations were discussed. It was found that the analytical solution produced accurate results, but reductions in computation time compared to numerical simulation were unable to be achieved. In particular, a major factor in the lengthy computation time was the fact that the calculation had to be run out to at least nine periods to get good results, because the current only converges past eight periods and multiple periods had to be summed together to get an average current.

INTRODUCTION

Strong field electron emission (SFEE) is the process in which electrons are emitted, such as from the surface of a metal, due to the application of a strong electric field. The intense electron emission induced by SFEE (*1*) can be used for many applications such as electron interferometry, nanometric imaging, and in synchrotrons (*2*). A schematic for an SFEE setup is shown in Figure 1. This setup is an example of laser-induced SFEE, where a laser beam provides the intense electric fields necessary to eject electrons from the surface. The motivation for this paper is the need to create and investigate yield curves describing laser-induced SFEE. A yield curve is a plot relating induced electron current (yield) to applied electric field magnitude. Yield curves are relevant to the applications of SFEE which are sensitive to the electron current because knowing the theoretical yield informs the necessary laser strength. Current methods of obtaining yield curves for SFEE involve complex time-intensive simulations which could be bypassed by directly evaluating the analytical solution to the 1-D time-dependent Schrödinger equation (TDSE). A numerical solution to the TDSE for strong laser field emission is tractable but takes extensive computation time. In the finite difference scheme, the spatial grid step size must be able to support the high energy electrons that are emitted, that is, $dx \ll \frac{1}{k_{max}} \sim \frac{1}{\sqrt{E_{max}}}$. Likewise, the time step must satisfy $dt \ll \frac{1}{\omega_{max}} \sim \frac{1}{E_{max}}$. The stability and accuracy requirements of the particular numerical method in use must also be satisfied. The spatial grid must extend far into the vacuum and far into the material to avoid reflections from the simulation boundaries or the metallic slab potential backing for the duration of the laser pulse. Continuous wave calculations require absorptive or transparent boundaries to avoid some of these issues. Additionally, a Heaviside step potential is highly sensitive to the spatial grid size – the increase in potential which nominally takes place over zero distance now takes place over a distance of dx, making studies of this ideal system less reliable. Due to these requirements, it was hypothesized that an analytical calculation would be more computationally efficient than performing numerical simulations. This hypothesis was investigated by plotting and comparing yield curves from both the analytical and numerical methods.

To increase electron yield in SFEE, a sharp geometry is used to enhance surface field intensity by concentrating field lines (*3*). The phenomenon of sharp geometry concentrating electric field lines is known as the lightning rod effect, shown in Figure 2. One commonly used structure for sharp geometry is the nanotip - a nanoscopic needle. The geometry of a nanotip can

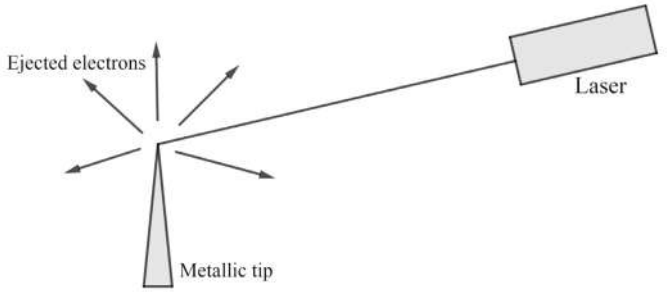


Figure 1. Experimental setup for laser-induced SFEE. The vertically polarized laser beam creates a strong, oscillating electric field at the surface of the metallic tip causes electrons to be ejected from the tip through strong field electron emission (SFEE). The sharp, needle-like shape of the tip enhances the surface electric field intensity through the lightning rod effect, increasing yield of the emitted electron current.

be approximated by a cone having a spherical tip with radius of curvature of ~ 10 nm. This paper uses a one-dimensional approximation of a nanotip.

The use of a 1-D approximation to represent the nanotip is justified by the following length scales. Firstly, the Fermi energy of gold (and therefore, the bound kinetic energy of the most easily excited electrons) is 5 eV. This corresponds to the Fermi wavelength on the scale of 0.1 nm, because $\frac{\hbar^2 k^2}{2m_e}$ = 5 eV which implies $\lambda \approx 0.6$ nm. Secondly, the mean transverse energy (MTE) of emitted electrons is ~0.6 eV. For laser field emission with an 800 nm laser, an electron of that energy travels 0.3 nm over $\frac{1}{4}$ of the laser cycle (0.7 fs). Lastly, the minimum spread of the wavefunction permitted by the Heisenberg uncertainty principle, over this same time interval, is ~0.3≪10 nm (*3*). All these quantities are much less than the nanotip radius of curvature of ~ 10 nm, and thus do not present issues for using a 1-D approximation to calculate nanotip yield scaling.

For laser-induced SFEE, there are two dominant electron emission mechanisms: multi-photon absorption at low fields, and quantum tunneling at high fields. In the low-field multi-photon regime, the emitted current is expected to follow a power law of the form $n_e \propto I^{n_\gamma}$ where I is the laser intensity which is proportional to E^2 and n_γ is the number of photons required to emit an electron ($n_\gamma = \lceil \frac{W}{\hbar\omega} \rceil$) (*4, 5*). At high fields where quantum tunneling is dominant, it is expected to see a tapering off in the yield scaling. The ponderomotive energy, which is the average quiver energy of the electron in an oscillating electric field, increases quadratically with the field strength as $U_p = \frac{E^2\lambda^2}{16\pi^2 c^2}$ in atomic units. Electron states in the laser field thus have an increased energy, effectively raising the vacuum potential by the ponderomotive energy (or ponderomotive potential). Thus, as the applied field strength increases, the ionization barrier also increases. This eventually increases the number of photons necessary for ionization and reduces current by channel closing, which results in noticeable dips in the yield curve (*6*). These characteristics should be reproduced by the yield curves.

MATERIALS AND METHODS

1-D approximation

In the 1-D model which provides a simple representation of the nanotip, the electron wavefunction evolves as

$$i\partial_t \psi(x,t) = -\frac{1}{2}\Delta\psi(x,t) + \Theta(x)(U - Ex\,\Theta(t)\cos\cos\omega t)\psi(x,t) \quad \text{(Eq. 1)}$$

where Θ is the Heaviside step function, U is the metallic potential depth (typically the work function added to the Fermi energy), and E is the magnitude of the applied laser field (*7*). In this paper, atomic units are used where $\hbar = m_e = 1$. This equation represents a scenario where there is a semi-infinite metallic slab from $x = -\infty$ to $x = 0$ and an oscillating electric field from $x = 0$ *to* $x = \infty$ that turns on at time $t = 0$. Before turning on the electric field, there exists the stationary solution

$$\varphi_0(x) = \{e^{ikx} + \frac{ik+\sqrt{2U-k^2}}{ik-\sqrt{2U-k^2}}e^{-ikx} \; for\; x < 0 \quad \frac{2ik}{ik-\sqrt{2U-k^2}}e^{-\sqrt{2U-k^2}x} \; for\; x > 0 \quad \text{(Eq. 2)}$$

which preserves continuity of ψ and $\partial_x\psi$ at $x = 0$. The probability current is calculated by

$$j(x,t) := I(\psi^*(x,t)\partial_x\psi(x,t)) \quad \text{(Eq. 3)}$$

where I is the imaginary part. The probability current at $x = 0$ represents the average movement of electrons at the surface of the metal which may be directly related to the charge current or yield.

Analytical solution

The solution to this equation at $x = 0$ as derived by Costin et al. is:

$$\psi(0,t) := \psi_0(t) = h(t) + L\psi_0(t) \quad \text{(Eq. 4)}$$

with $h(t)$ defined by:

$$h(t) = h_+(t) + h_-(t) - \frac{1}{\pi}\int_0^t du h_-(u)\int_u^t ds\frac{g(s,t)}{\sqrt{s-u}} \quad \text{(Eq. 5)}$$

$$h_-(t) = \frac{1}{2}e^{-i\frac{k^2}{2}t}\left[erfc\left(-e^{-i\frac{\pi}{4}}\sqrt{\frac{t}{2}}k\right) + Rerfc\left(e^{-i\frac{\pi}{4}}\sqrt{\frac{t}{2}}k\right)\right] \quad \text{(Eq. 6)}$$

$$h_+(t) = Texp\left[\frac{E}{\omega^2}D(1 - \cos\cos\omega t) - i\left(\frac{k^2}{2} + \frac{E^2}{4\omega^2}\right)t + i\frac{E^2}{8\omega^3}\sin\sin 2\omega t\right] \times erfc\left[e^{-i\frac{\pi}{4}}\left(i\sqrt{\frac{t}{2}}D + \frac{E}{\omega^2}\frac{1-\cos\cos\omega t}{\sqrt{2t}}\right)\right] \quad \text{(Eq. 7)}$$

and the integral operator L is defined by:

$$L\psi_0(t) = \frac{E}{2\omega\sqrt{2i\pi}}\int_0^t ds\psi_0(s)\frac{\alpha(s,t)}{\sqrt{t-s}}e^{if(s,t)} + \frac{1}{2\pi}\int_0^t du\psi_0(u)\int_u^t ds\frac{g(s,t)}{\sqrt{s-u}} \quad \text{(Eq. 8)}$$

with constants reminiscent of the reflection and transmission coefficients:

$$D = \sqrt{2U - k^2}\; R = \frac{ik+D}{ik-D}\; T = \frac{ik}{ik-D} \quad \text{(Eq. 9)}$$

and miscellaneous functions:

$$f(s,t) = \frac{E^2(\cos\cos\omega t - \cos\cos\omega s)^2}{2\omega^4(t-s)} - \left(U + \frac{E^2}{4\omega^2}\right)(t-s) + \frac{E^2}{8\omega^3}(\sin\sin 2\omega t - \sin\sin 2\omega s) \quad \text{(Eq. 10)}$$

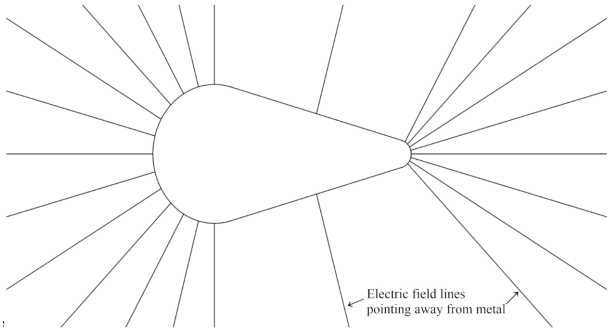


Figure 2. The lightning rod effect. A charged conductor emanating electric field lines with uniform angular sA charged conductor emanating electric field lines with uniform angular spacing. The areas with sharper geometry have a higher density of field lines and therefore a higher surface electric field intensity, while the flat surfaces have almost no field lines. This demonstrates how sharp geometry enhances surface electric field intensity.

$$g(s,t) = \frac{e^{if(s,t)}-1}{2(t-s)^{\frac{3}{2}}} + i\frac{\partial_s f(s,t)e^{if(s,t)}}{\sqrt{t-s}} \qquad \text{(Eq. 11)}$$

$$\alpha(s,t) = \sin sin\,\omega s + \frac{coscos\,\omega t - coscos\,\omega s}{\omega(t-s)} \qquad \text{(Eq. 12)}$$

$$\partial_s f(s,t) = \frac{E^2}{4\omega^4}\frac{(coscos\,\omega t - coscos\,\omega s)(coscos\,\omega t - coscos\,\omega s + 2\omega(t-s)sinsin\,\omega s)}{(t-s)^2} +$$

$$U + \frac{E^2}{4\omega^2}(1 - \cos cos\,2\omega s) \qquad \text{(Eq. 13)}$$

The spatial derivative necessary to obtain the current is given by

$$\partial_x \psi_0(t) = \sqrt{\frac{2}{i\pi}}\partial_t\left[\psi_0(t) \star t^{-1/2} - 2h_-(t) \star t^{-1/2}\right] \qquad \text{(Eq. 14)}$$

with ★ being the Laplace convolution

$$[f \star g](t) = \int_0^t ds f(s) g(t-s) \quad \text{(Eq. 15)}$$

(*7*). The variables used, field strength E, wave vector in the metal k (typically $k^2/2 = E_f$), laser frequency ω, and metallic potential depth U (typically $U = E_f + W$), are in atomic units. Initial wave function is $\psi_0(0) = 2T$ which matches Costin et al. (*7*). In this paper, the physical values used are $\omega = 1.55\,eV = 0.0570\,a.u.$, $U = 10\,eV = 0.3676\,a.u.$, and $k = \sqrt{2E_f} = 0.5752\,a.u.$, which match both Yalunin et al. and Costin et al. (*6, 7*).

Chebyshev decomposition

As suggested by Costin et al. (*7*), ψ_0 is decomposed into a Chebyshev polynomial series. The function is split into segments of length $\Delta t = 2\pi/\omega$ (the duration of a full laser cycle) to accommodate the asymptotic behavior which is $2\pi/\omega$ periodic. To account for the singular behavior of $\psi_0(t) \propto t^{\frac{3}{2}}$ near $t = 0$, an additional term is included in the decomposition,

$$\psi_0^{(l)}(t) \approx c_{-1}^{(l)} T_{-1}(t) + \sum_{j=0}^{N} c_j^{(l)} T_j\left(\frac{\omega t - (2l+1)\pi}{\pi}\right) \qquad \text{(Eq. 16)}$$

where segment number $l = \lfloor \omega t/2\pi \rfloor$, $T_{-1}(t) = t^{\frac{3}{2}}$, and T_0 through T_N are Chebyshev polynomials of the first kind, evaluated on the domain [-1,1). The integral equation can be rewritten as

$$\psi_0^{(l)}(t) = h(t) - \int_{2l\pi/\omega}^{t} ds m(s,t)\psi_0^{(l)}(s) - K_l(t) \qquad \text{(Eq. 17)}$$

where $m(s,t)$ is a function independent of $\psi_0(t)$:

$$m(s,t) = -\frac{E}{2\omega\sqrt{2i\pi}}\frac{\alpha(s,t)}{\sqrt{t-s}}e^{if(s,t)} - \frac{1}{2\pi}\int_s^t du \frac{g(u,t)}{\sqrt{u-s}} \qquad \text{(Eq. 18)}$$

and $K_l(t)$ is the integral of the already calculated wavefunctions

$$K_l(t) = \sum_{p=0}^{l-1}\int_{2p\pi/\omega}^{2(p+1)\pi/\omega} ds m(s,t)\psi_0^{(p)}(s) \quad \text{(Eq. 19)}$$

with $K_0(t) = 0$. Now

$$c_{-1}^{(l)}\left[T_{-1}(t) + I_{-1}(t)\right] + \sum_{j=0}^{N} c_j^{(l)}\left[T_j\left(\frac{\omega t-(2l+1)\pi}{\pi}\right) + I_j(t)\right] = h(t) - K_l(t) \quad \text{(Eq. 20)}$$

where $I_j(t)$ is

$$I_j(t) = \int_{2l\pi/\omega}^{t} ds m(s,t) T_j\left(\frac{\omega s-(2l+1)\pi}{\pi}\right) \qquad \text{(Eq. 21)}$$

for $j \neq -1$ and $\int_{2l\pi/\omega}^{t} ds m(s,t) T_{-1}(t)$ for $j = -1$. To find the $N+1$ unknown $c_j^{(l)}$'s, $N+1$ values of t can be substituted into equation (20) and the resulting system of $N+1$ equations can be solved. That is, the matrix equation $A_{li}^{(l)} c_i^{(l)} = b_i^{(l)}$ yields the desired solution, where

$$A_{ij}^{(l)} = T_j\left(\frac{\omega t_i-(2l+1)\pi}{\pi}\right) + I_j(t_i)\, b_i^{(l)} = h(t_i) - K_l(t_i) \quad \text{(Eq. 22)}$$

for $j \neq -1$ and $A_{ij}^{(l)} = T_{-1}(t_i) + I_{-1}(t_i)$ for $j = -1$. The t_i are chosen to be distributed throughout the segment under consideration. It was found that better results are obtained when t_i is sampled according to the following function which has an increased sample density near the boundaries of each segment,

$$t_i = \frac{\pi}{\omega}\left[\left(1 + \cos cos\,\frac{\pi i}{N}\right) + 2l\right], \quad t_i \in \{0, 1, \ldots, N\} \qquad \text{(Eq. 23)}$$

Numerical singularities, series expansions, and integration

Some of these functions contain indeterminate forms, singularities, or both, which cause numerical errors as $t \to s$. To resolve these errors, Taylor series approximations are used when $|\omega(t-s)| < 10^{-6}$. Singular integrals are calculated numerically using Chebyshev-Gauss quadrature and the Laplace convolution is evaluated using Gauss-Jacobi quadrature.

Frequency limit of chebyshev polynomials

As the applied electric field amplitude increases the number of polynomials n must be increased. The Chebyshev polynomial decomposition uses elements of the form $T_n(\frac{\omega_g}{\pi}t) = T_n(t')$. A Chebyshev polynomial of order n effectively undergoes an $n\pi$ phase shift on the region [-1,1]. Therefore, the maximum angular frequency allowed is $\omega'_{max} = \frac{n\pi}{2} \Rightarrow \omega_{max} = \frac{\omega_0}{\pi}\omega'_{max} = \frac{n}{2}\omega_0$. Consequently, n should be set such that $n > 2\frac{\omega_{max}}{\omega_0}$ in order to ensure the resolution of the highest frequencies. Semi-classically, the

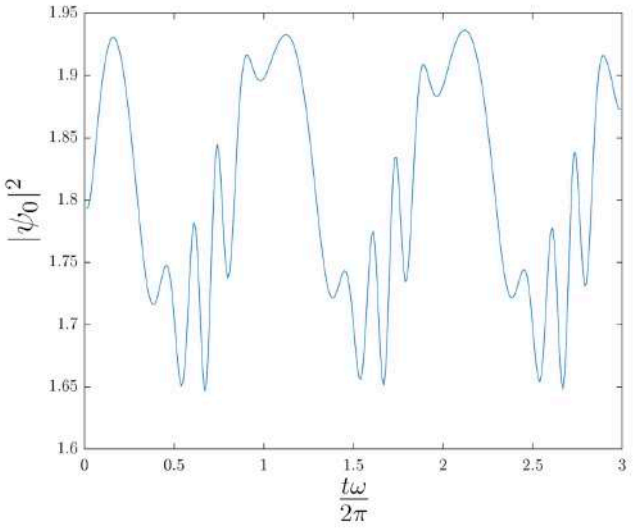


Figure 3. $|\psi_0|^2$ vs time for the first three periods. Surface probability density $|\psi_0|^2$ over time of an electron in metal subject to three cycles of an oscillating laser field computed using the analytical solution derived by Costin et al. (7) with E=15 V/nm. Figure 3 is nearly identical to Figure 1 in Costin et al. (7), confirming proper implementation of the analytical code. Similar results are obtained from numerical simulation, so the analytical and numerical methods of solving the TDSE for the electron's wavefunction agree.

highest possible electron kinetic energy is $T_{max} = \omega_{max} = 10U_p = \frac{5}{2}\frac{E^2}{\omega_0^2}$ in atomic units (*3*). Substituting the value for ω_{max} yields $n > 5\frac{E^2}{\omega_0^3}$. A reasonable overshooting by 50% to ensure all necessary frequencies are resolved would be $n \approx 7.5\frac{E^2}{\omega_0^3}$. It was found that better results are obtained when setting a minimum n of around 100 even at low electric field amplitudes, likely necessary due to the frequency-rich singularity at $t = 0$.

Numerically solving the TDSE

Equation 1 can also be solved with numerical methods such as the split-step and Crank-Nicolson methods. To do this, the semi-infinite slab is approximated as a finite potential well with depth - $W - E_f$. The width of the well should support a bound eigenstate at the Fermi level - W, which would be set as the initial wavefunction. A buffer is maintained on each side of the well to reduce boundary effects and let the wavefunction evolve according to the TDSE. The results from the numerical methods are used to verify the analytical calculations.

Determining appropriate time boundaries

After some initial attempts, it became apparent that the calculation would have to be run for a longer time period before the current would converge, so the creation of Figure 5 was necessitated. Figure 5 confirmed that only having the early periods was insufficient and indicated that calculations would

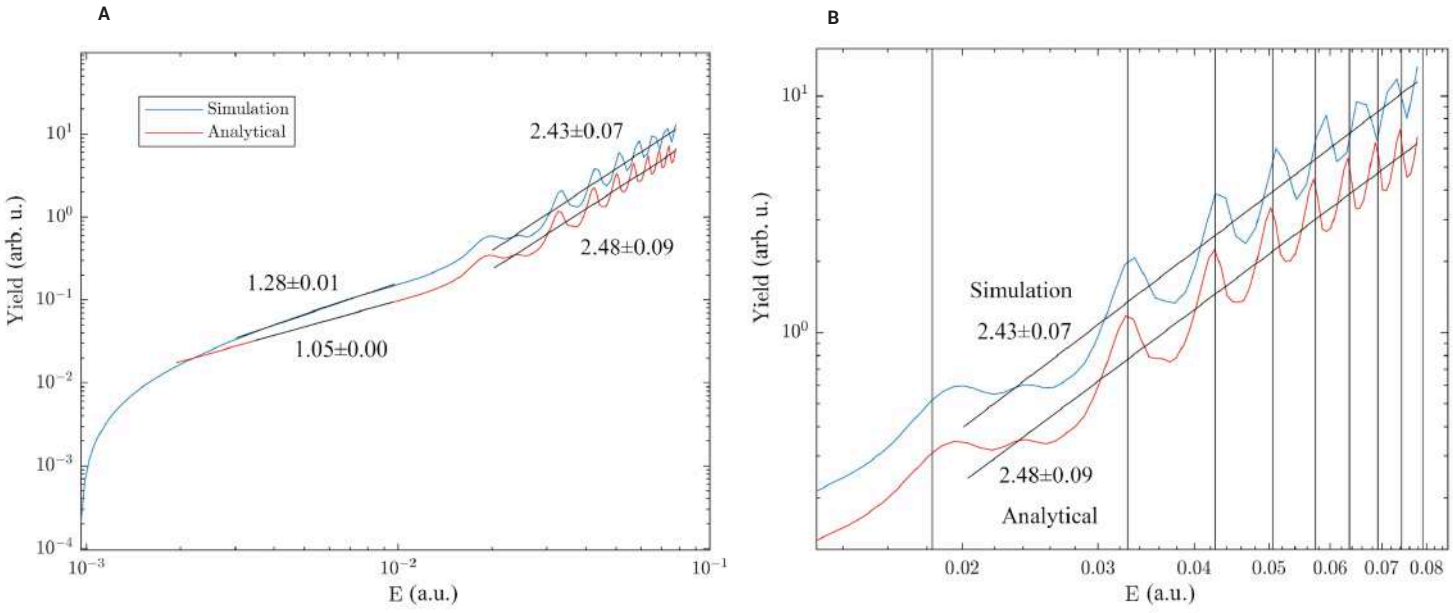


Figure 4A and 4B. Periods 1-4 yield curve. Electron probability current leaving the metal surface integrated over periods 1-4 plotted against the. electric field amplitude of the applied oscillating laser field, computed using the analytical solution derived by Costin et al. (7). Figure 4B is an enlarged portion of Figure 4A, with added vertical lines at the locations where the sum of work function and the ponderomotive energy is an integer multiple of photon energy, representing where channel closing occurs. The fit and error parameters are the slope and standard deviation for a least squares linear fitting, representing the power law of yield scaling on a log-log plot.

have to be run past around eight periods to accurately reflect the long-term current. $\psi(x,t)$ is proven to converge to its asymptotic behavior as $t^{\frac{-3}{2}}$ (7).

RESULTS

Reproduction of previous findings

The output of the analytical calculation for the surface probability density $|\psi_0|^2$ over time with E = 15 V/nm = 0.0292 a.u. is graphed in Figure 3. The result agrees well with both Costin et al.'s results (7) and the numerical results, indicating that the analytical code was implemented correctly.

Yield curve periods 1-4

The first yield curve, the integrated probability current over the first four periods as a function of the applied field strength E, is displayed in Figure 4. The wave function was not normalized for the simulation, so the vertical axis units could be different for the simulation and analytical curves. As expected, there are dips in yield where channel closing occurs. However, while eighth power yield scaling for four-photon emission is expected, the yield scaling is actually approximately linear.

Rate of convergence of this setup

To determine how long to run the calculation for, current against the period for electric field amplitudes of 1, 3, 15, 30, and 40 V/nm are plotted using numerical simulation. The 30 V/nm curve, shown in Figure 5, converged the slowest, after eight periods.

Yield curve periods 6-9

The second yield curve, shown in linear and log-log plots in Figure 6, integrates the total current from periods 6-9 using the analytical calculation and plots it versus the electric field amplitude. This curve terminates right as the four-photon channel closes, so the channel closing oscillations are not shown. Otherwise, the expected eighth power yield scaling is seen. The average current of each period is plotted individually in Figure 7. Each of the single-period yield curves oscillates a few times, and at some points the oscillations go deep enough into the negative currents to make the total current negative, but these oscillations are mitigated by summing together multiple periods.

DISCUSSION

In this paper it was attempted to generate yield curves modeling SFEE by numerically evaluating the analytical solution to the TDSE derived by Costin et al. (7). First, the validity of the analytical code needed to be ensured by plotting the surface probability density in Figure 3 and confirming its agreement with Costin et al.'s results (7) and with the numerical results. The first yield curve, Figure 4, demonstrates channel closing. Once the ponderomotive energy eliminates a photon channel, the yield drops off sharply before the next channel picks up to dominate emission. The unexpected yield scaling has two explanations. To begin with, the first period's current is anomalous, in that it is much greater than

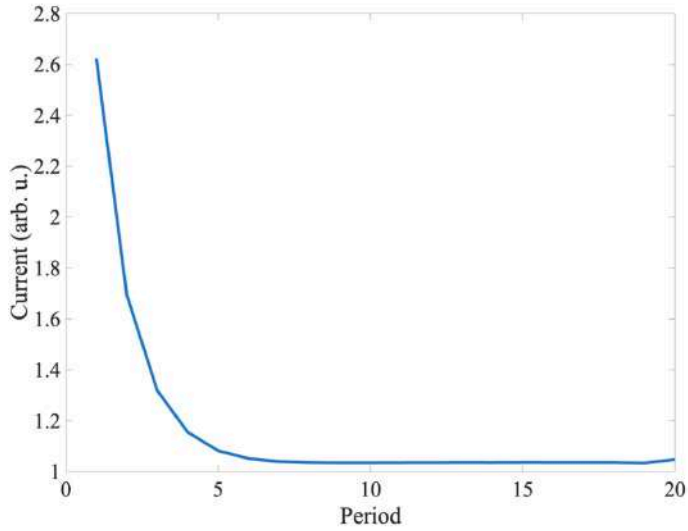


Figure 5. Current vs period for E=30 V/nm. Electron probability current leaving the metal surface due to the applied oscillating laser field with amplitude having slowest converging current (E=30V/nm). Plotted against period, computed using numerical simulation. This demonstrates convergence of current over time, which is why the first few oscillations should be discarded for yield scaling calculations.

all the other period's currents and scales approximately linearly. This could be due to the unphysical instantaneous and therefore discontinuous application of the electric field at $t = 0$ from the cosine term in (Eq. 1). This discontinuity is frequency-rich with no temporal scale and therefore no frequency scale. Thus, the multi-photon absorption picture cannot apply and the yield attributed to this feature is non-trivial. Secondly, the current has not yet converged to the steady state within the first four periods that are under consideration. Short pulse simulations attain the proper yield scaling (6, 8) before approaching a steady-state solution, making the former point (the anomalous first period current) the likely explanation.

The second yield curve, Figure 6, is taken from later periods and demonstrates the expected eighth power yield scaling, which is corroborated by Figure 4 in Yalunin et al. (6). The maximum electric field amplitude of Figure 6 could not be extended past 10 V/nm due to computational time constraints. Additionally, many of the integrated current data points result in a negative total yield. This is due to the transient oscillations in the probability current and would not be observed in the steady state. This effect can be visualized by plotting the average current of each period individually, shown in Figure 7.

Yield curves that exhibited the correct power law scaling were successfully produced, and other yield curves that exhibited dips as photon channels closed were also able to be produced. Unfortunately, due to computational time constraints, a yield curve that exhibited both at once proved unfeasible to produce. In general, the analytical and numerical calculations gave similar results, but the numerical calculations finished much faster (hours vs weeks) than the analytical ones as the parameters

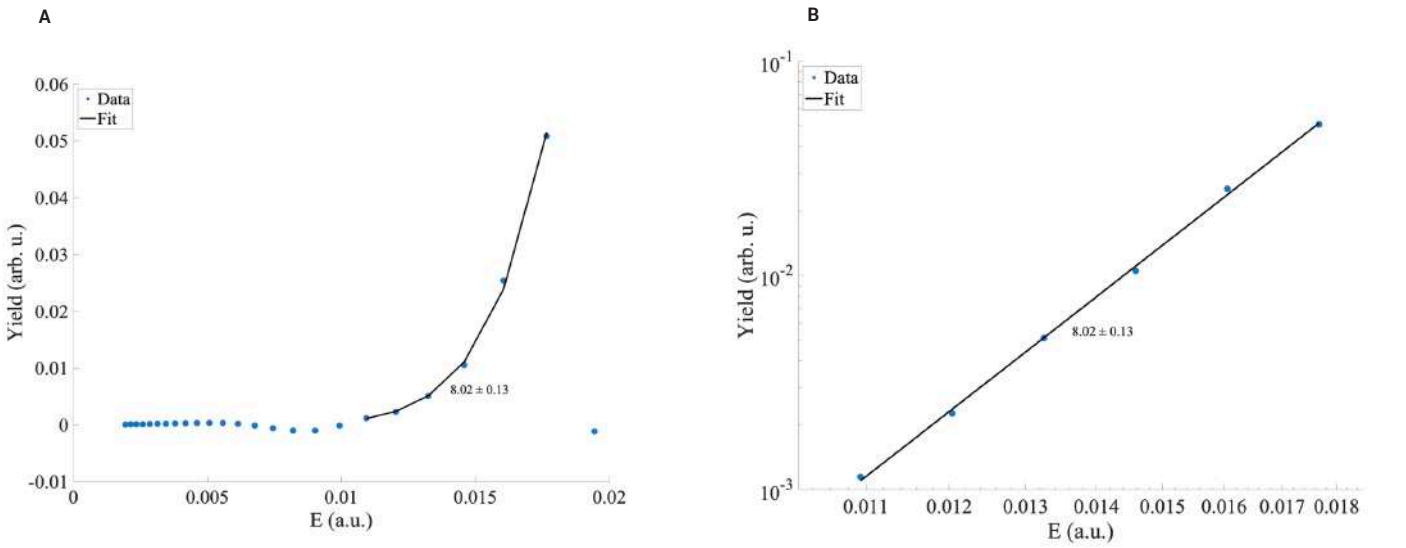


Figure 6A and 6B. Periods 6-9 yield curve. Electron probability current leaving the metal surface integrated over periods 6-9 plotted against electric field amplitude of an applied oscillating laser field, computed using the analytical solution derived by Costin et al. (7). The fit and error parameters are the slope and standard deviation for a least squares linear fitting on the log-log plot in Figure 6B, which on Figure 6B is a power law of the yield scaling. The curve terminates as the four-photon channel closes, so channel closing oscillations are not observed.

entered the region of larger time intervals, higher electric field amplitudes, or both.

One reason why the numerical calculations finished considerably faster than the analytical ones is that the numerical simulation has lower order computational scaling. For the analytical calculation, for each period, integrals had to be calculated (including unique integrals of previous period results) proportional to the N^2 where N is the number of polynomials used in the decomposition. This was followed by a relatively cheap $(O(N^3))$ matrix inversion. Then, the forward calculation to attain the wave function values required $\propto N^2$ evaluations of Chebyshev polynomials for each time step sample. Next, calculating the Laplace convolutions of the spatial derivative in order to find the surface current took days, a similar amount of time as all the above steps combined. Additionally, the number of integral steps had to be increased as the period number increased because $\psi_0(t)$ depends on an integral from 0 to t, adding an additional factor to its computational time scaling. This time scaling difficulty is compounded by the cosine discontinuity at $t = 0$ in (Eq. 1) which means that the first few periods give unphysical results. This is a good reason to prefer steady-state results for yield calculation, such as Costin et al.'s other paper that considers the limit as time approaches infinity (*9*).

Additionally, the number of Chebyshev polynomials used had to be increased as electric field amplitude increased in order to resolve the higher frequencies. A calculation was run with high electric field amplitudes which demonstrated channel closing (Figure 4) and another calculation with longer time boundaries which demonstrated the correct power law scaling (Figure 6) was also run. However, a calculation with both high electric field amplitudes and longer time boundaries that could demonstrate both channel closing and the correct power law scaling at once would take on the order of weeks to run whereas a numerical approach would take hours. Nevertheless, Costin et al.'s analytical solution (*7*) produced accurate results (graphs that were visually identical to those produced using numerical simulation), so with sufficient optimization and computational resources, this solution is viable to generate yield curves for SFEE. A possible method of reducing computation time would be to change the evaluation methods, such as by using Fourier decomposition instead of Chebyshev, allowing for the use of fast Fourier transforms. For future research, one possibility is to run more yield curve calculations using an analytical solution to a pulsed system, such as the one derived in Luo et al. (*10*). A spectral filter could be applied to model image charge trapping (*8*) and estimate various beam parameters including the mean transverse energy (MTE) which is of import for most applications. MTE is important because it is related to emittance, a quantity describing a beam's size in phase space and which is oftentimes conserved in accelerator physics in the absence of space-charge. The emittance is further related to brightness, which describes a beam's density in phase space. Beams that are brighter have more electrons, a smaller spatial size, and/or a smaller momentum spread (corresponding to smaller MTE). There are multiple kinds of brightness used for various applications, but all of them relate to the MTE somehow. For X-ray free electron lasers, the achievable photon brightness depends on the beam brightness (11). In the case of ultrafast transmission electron microscopes, getting the best resolution possible requires certain beam emittance thresholds to be met (12). By knowing the MTE

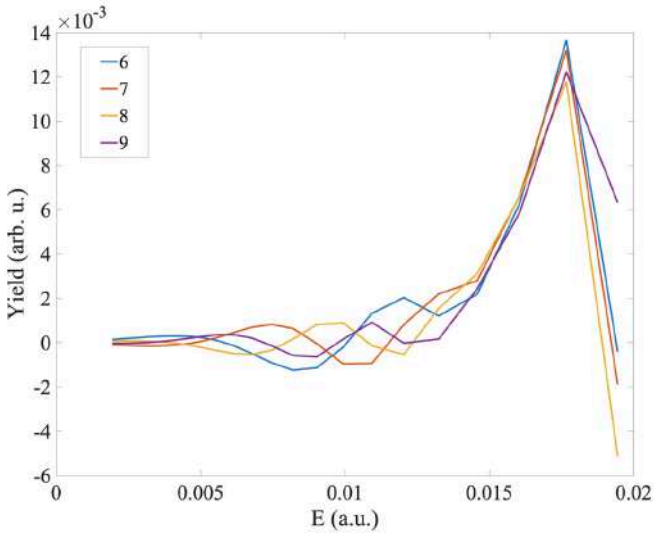


Figure 7. Yield curves for periods 6, 7, 8, and 9. Electron probability current leaving the metal surface for individual periods 6, 7, 8, and 9 (the constituent periods forming Figure 6) vs. electric field amplitude of an applied oscillating laser field, computed using the analytical solution derived by Costin et al. (7). The figure depicts transient oscillations in the probability current that cause some of the points in Figure 6 to have negative current, and further, that integrating current over more periods smooths out these oscillations.

of a laser field emission source, the emittance and brightness of the emitted electron beams could be calculated, and therefore the source's efficacy for these applications and more could be predicted. These pulsed system studies may be compared to ab-initio models of a similar system which include a Jellium potential, a full metallic electron distribution, image charge trapping, and dispersion relation transitions. Because such models take days to weeks to evaluate, there is a lot of opportunity for improvement in terms of time savings if similar results can be achieved by a faster calculation.

ACKNOWLEDGEMENTS

This research was funded by the Center for Bright Beams, National Science Foundation Grant No. PHY-1549132. Big thank you to Joshua Mann and Professor James B. Rosenzweig for making this research possible.

REFERENCES

1. C. Ropers, D. R. Solli, C. P. Schulz, C. Lienau, T. Elsaesser, Localized multiphoton emission of femtosecond electron pulses from metal nanotips. *Phys. Rev. Lett.* **98**, 043907 (2007). doi:10.1103/PhysRevLett.98.043907
2. P. Hommelhoff, Y. Sortais, A. Aghajani-Talesh, M. A. Kasevich, Field emission tip as a nanometer source of free electron femtosecond pulses. *Phys. Rev. Lett.* **96**, 077401 (2006). doi:10.1103/PhysRevLett.96.077401
3. M. Krüger, C. Lemell, G. Wachter, J. Burgdörfer, P. Hommelhoff, Attosecond physics phenomena at nanometric tips. *J. Phys. B: At. Mol. Opt. Phys.* **51**, 172001 (2018). doi:10.1088/1361-6455/aac6ac
4. J. Bechtel, W. Smith, N. Bloembergen, Four-photon photoemission from tungsten. *Opt. Commun.* **13**, 56–59 (1975). doi:10.1016/0030-4018(75)90156-X
5. G. Ferrini, F. Banfi, C. Giannetti, F. Parmigiani, Non-linear electron photoemission from metal with ultrashort pulses. *Nucl. Instrum. Meth. A.* **601**, 123–131 (2009). doi:10.1016/j.nima.2008.12.107
6. S. Yalunin, M. Gulde, C. Ropers, Strong-field photoemission from surfaces: Theoretical approaches. *Phys. Rev. B.* **84** 195426 (2011). doi:10.1103/PhysRevB.84.195426
7. O. Costin, R. Costin, I. Jauslin, J. L. Lebowitz, Exact solution of the 1D time-dependent Schrödinger equation for the emission of quasi-free electrons from a flat metal surface by a laser. *J. Phys. A: Math. Theor.* **53** 365201 (2020). doi:10.1088/17518121/aba1b6
8. J. Mann, T. Arias, G. Lawler, J. Nangoi, J. Rosenzweig, Simulations of nanoblade-enhanced laser-induced cathode emissions and analyses of yield, MTE, and brightness. Presented at the 12th International Particle Accelerator Conference, Campinas, SP, Brazil, 1 September 2021, 2957–2960. doi:10.18429/JACoW-IPAC2021-WEPAB147
9. O. Costin, R. Costin, I. Jauslin, J. L. Lebowitz, Solution of the time dependent Schrödinger equation leading to Fowler-Nordheim field emission. *J. Appl. Phys.* **124**, 213104 (2018). doi:10.1063/1.5066240
10. Y. Luo, Y. Zhou, P. Zhang, Few-cycle optical-field-induced photoemission from biased surfaces: An exact quantum theory. *Phys. Rev. B.* **103**, 085410 (2021). doi:10.1103/PhysRevB.103.085410
11. P. Emma et al., First lasing and operation of an ångstrom-wavelength free-electron laser. *Nat. Photonics.* **4**, 641–647 (2010). doi:10.1038/nphoton.2010.176
12. D. Xiang, F. Fu, J. Zhang, X. Huang, L. Wang, X. Wang, W. Wan, Accelerator-based single-shot ultrafast transmission electron microscope with picosecond temporal resolution and nanometer spatial resolution. *Nucl. Instrum. Meth. A.* **759**, 74-82 (2014). doi:10.1016/j.nima.2014.05.068.

2021 - 2022

UCLA USJ STAFF