Carbon Currencies: Isotopic constraints on the biogeochemistry of organic acids

Thesis by
Elliott P. Mueller

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy in Geobiology

Caltech

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2024
Defended May 3, 2024

© 2024

Elliott P. Mueller
ORCID: 0000-0002-6837-0409

When the microscope is improved,
we shall have the cells analyzed,
and all will be electricity
of somewhat else.

- Annie Dillard
Pilgrim at Tinker Creek (1974)

ACKNOWLEDGEMENTS

It is hard to capture my gratitude for my time at Caltech. When I first arrived, my scientific interests were difficult to pin down. I tended to flit about, landing wherever was most enticing. It was at Caltech that I learned how to harness this curiosity. In fact, during graduate school, I learned how to integrate many sides of my personality into my work as a scientist. These lessons made me a better student, a more engaged mentor, a more empathetic collaborator and a more thoughtful experimentalist. I am grateful to the faculty that supported me throughout this process: Alex Sessions, John Eiler, Jared Leadbetter, Woody Fischer, Victoria Orphan, and Dianne Newman. Alex, your rigorous and thoughtful approach to science will always be a source of inspiration for me. You recruited me thinking I would do some geochemistry and I lost myself in the world of microbiology. Thank you for letting me find my own way through this thesis, even when it didn't seem to be the most prudent path. Your faith in me made my research possible. It never went unnoticed. John, I have learned so much from working with you. Thank you for helping me grow into an independent scientist, for always being open to debate, and for letting me run around your lab for five years. Woody, your passion for geobiology is contagious. It made me feel welcome in a community that, at first felt foreign. Jared, your class (and later Dianne's) were the first microbiology classes I ever took. It's not an understatement to say that they changed the way I see the world, and it was so gratifying to teach that same class with you only three years later. I am also grateful to Victoria Orphan, who made time to see me, let me use her various pieces of equipment, and brought me into a collaboration that became Chapter V of my thesis. I had lots of support from mentors outside of Caltech, people who changed my view points and encouraged me to keep pursuing my research. Kai-Uwe Hinrichs, Barbara Sherwood Lollar, Matthias Hess, Reto Wijker, Nagissa Mahmoudi, Ludmilla Aristilde, Andreas Hilkert, Kate Freeman and Roger Summons to name a few. Each delivered their advice with kindness and understanding.

One of my favorite parts of doing a Geobiology Ph.D. at Caltech was the endless list of instruments and equipment at our disposal. However, these resources are nothing without (or in comparison to) the technical expertise of laboratory managers and technicians in the Geological and Planetary Sciences division. Fenfang Wu, Nathan Dalleska, and Stephanie Connon, in particular, were invaluable for my work. I

could go on about either their intellect or organizational prowess. What I learned best from you all though is that no problem is insurmountable. Often, I felt alone solving daily problems in the lab. It can be an overwhelming feeling. Fenfang, Nathan, and Stephanie showed me that cool heads prevail.

My time at Caltech was punctuated by several stints at other research institutions. Those experiences shaped me as much as my time at Caltech. Thank you to Peter Willis and Jessica Creamer at NASA JPL for taking me in as a first-year. At the end of the third year, I went to the Microbial Diversity course in Woods Hole, MA. If I wasn't already hooked on microbial sciences, this sealed the deal. It was an idyllic summer filled with incredible friends and colleagues. I especially want to think George O'Toole and Rachel Whitaker for showing me new ways of approaching science. Finally, in my fourth year, I spent a semester in Bremen, Germany working in Kai-Uwe Hinrichs group at MARUM and in Andreas Hilkert's group at Thermo Fisher. There I got to work with Verena Heuer, who continues to be one of the most rigorous analytical scientists I have ever met. Her work ethic is matched by her kindness and generosity. That experience also introduced me to new friends that I am grateful for like Kristýna Kantnerova and Nils Kuhlbush.

I would be remiss to ignore the musicians that inspired me during my graduate work, particularly in moments of doubt. Artists like Big Thief, Arlo Parks, Jungle, Indigo De Souza, Babe Rainbow, Remi Wolf, Sidney Gish, Still Woozy, KAYTRANADA, Rostam, Adrianne Lenker, HAIM, Waxahatchee, DRAMA, Frank Ocean, Goth Babe, Palehound, Avett Brothers, Belle and Sebastian, and Generationals have been my constant companions for six years. They don't know it, but I am a better story teller because of them.

The friendships I've made in California have sustained me throughout graduate school. There are too many to list, but I'll try! Xenia Boyes, Sergio Parra, Makayla Betts, Shaelyn Silverman, Juliet Ryan-Davis, Sarah Zeichner, Philip Woods, and Emily Miaou (or the "Pit") you filled our first year with great memories. I was lucky enough to have Shae, Mak, Selva Marroquin, Alex Phillips, Hannah Dion-Kirschner, Eryn Eitel, Elise Wilkes, Juli Panehal and Yeonsoo Park as fellow Sessions Lab members during my time there. That community was always a source of support for me. I was also surrounded by a group of talented and friendly people in GPS that made every day at work interesting: Guannan Dong, Noam Lotem, Tim Csernica, Elle Chimiak, Amy Hofmann, Renee Wang, Rodney Tollerson, Madison Dunitz, Gabriella Weiss, Surjyendu Bhattacharjee, Simon Andren, James Mullahoo, Korbi

Thalhammer, and Lubna Shawar. I'm also grateful to have become close friends with a series of great roommates: Matthew Levine, Max Saccone, Charles Guan, and Ben Strozewky. I'd like to especially acknowledge Serg Parra for his friendship and many, many coffees and lunches together. Thank you for being your unwavering self.

While in Los Angeles, I maintained and built new friendships with people outside Caltech. Kelly Subin, Gabbie Gargano, Sarah Sprole, Jolie Bartner, Brennan Caruthers, Reid Flynn, Sean Geber, Anna Moreno-Takegami, Evan Gosselin, Caroline Hegg, Blaise Berglund, Audrey Spina, Arielle Sherman, Lilly Birk, Lindsey Scott, and Meghan Nosal, thank you for endless conversation, for never having any idea what I was doing, for visiting me from across the country, and for picking up the phone. I don't think I could have gotten through the Ph.D. without you.

To my partner, Joris: You have always supported me and believed in me. Thank you for filling our days with travel, friends, music and delicious food. Life feels light with you, and it's wonderful.

Finally, I am grateful to my family: My sisters, Kaitlyn and Colleen (and Tempie), and my parents, John and Eileen Mueller. I love you all very much. Kaitlyn and Colleen, your support over the past five years has meant the world to me. Dad, I have always been and continue to be inspired by your tenacity, your compassion, and your generosity. Thank you for teaching me to never settle and for building the road that led me here. Mom, you more than anyone have been there for me during this entire process. You effortlessly combine empathy, intelligence, and patience into an unstoppable force of good. There's never been a better role model, and your influence is woven throughout this thesis. Thank you.

ABSTRACT

On both human and geologic timescales, the microbial degradation of organic carbon in anoxic environments significantly influences the Earth's climate. The rate-limiting step of this process is the initial breakdown of complex organic polymers (e.g. cellulose) into small organic acids (e.g. acetate), which are then rapidly converted into either CO_2 or methane. While the steady-state concentration of organic acids is kept low by microbial turnover, the flux of reactions producing and consuming them is large. In my doctoral work, I leveraged this dynamic pool of metabolites as a window into the broader carbon cycle. Specifically, I developed novel analytical and computational tools that quantify and interpret the isotope composition of organic acids. These techniques provide new information about the mechanism and rates of organic acid turnover in nature.

First, in Chapter 2, I adapted electrospray ionization (ESI) Orbitrap mass spectrometry (MS) to simultaneously measure the carbon ($\delta^{13}C$) and hydrogen (δ^2H) isotope compositions of acetate. This approach is 50- to 1000-fold more sensitive than established techniques, making measurements of environmental samples feasible for the first time. This technique clearly distinguishes the metabolic sources of acetate (fermentation and acetogenesis). In Chapter 3, I developed a complementary computational tool to interpret this new isotopic information. Quantifying Isotopologue Reaction Networks (QIRN) builds numerical models of complex reaction networks, including metabolic pathways, and predicts the isotope composition of molecules produced by these networks. In Chapter 4, I combined my analytical and computational approaches to investigate the isotopic fractionations of the microbial metabolism that generate organic acids in nature, fermentation. I found that fermentation imposes a significant isotopic fractionation during the degradation of organic matter. By coupling flux-balance analysis and QIRN, I isolated the enzymes responsible for these fractionations. These results suggested that fermentation may have imprinted a carbon isotope trophic enrichment that is observable in the compound-specific $\delta^{13}C$ distribution of Proterozoic biomarkers. In Chapter 5, I used my Orbitrap method to quantify *in situ* acetate turnover rates based on the exchange of hydrogen atoms between water and acetate's methyl group. I took this tool to the environment, where I studied the biogeochemical drivers of carbon cycling in the deep continental subsurface. In Kidd Creek mine, which has subsurface fracture fluids that have been isolated for over a billion years, I found that acetate is being

actively produced and consumed in the subsurface. My analyses of acetate's isotope composition suggested that turnover may be driven by low-temperature water-rock reactions with implications for the habitability of subsurface environments elsewhere in the Solar System. Chapter 6 is a second application of the Orbitrap and QIRN in natural systems. This time I expanded the Orbitrap technique to include not just acetate but the C_3 and C_4 organic acids, propionate and butyrate. I investigated carbon turnover in the rumen fluid of cows, where microbial fermentation breaks down cellulose and transfers organic acids to the animal host. I found clear trends in the $\delta^{13}C$ and δ^2H of acetate and propionate that may hold information about the metabolic strategies of fermenters in the rumen. Finally, in Chapter 7, I highlight the challenges and opportunities of transitioning Orbitrap MS isotopic applications from pure standards to complex samples. These studies demonstrate bespoke strategies for isolating organic acids, and possibly other ESI-Orbitrap analytes, from environmental samples without fractionating their isotope ratios. Together, these chapters use a combination of novel analytical and computational tools to study the rate and mechanism of organic acid cycling in nature. Elucidating these drivers is necessary to understand the modern and ancient carbon cycle and to predict its response to climate change.

PUBLISHED CONTENT AND CONTRIBUTIONS

Mueller, Elliott et al. (2024). "Isotope ratio analysis of acetate using Orbitrap Exploris Isotope Solutions". In: *Thermo Fisher Technical Note* Isotope ratio MS.

E.P.M. conceptualized the project, performed the experiments, analyzed all samples, and wrote the manuscript.

Mueller, Elliott P., Juliann Panehal, et al. (2024). "Isotopic evidence of acetate turnover in the Precambrian continental crust". In: *Research Square Preprint*. DOI: 10.21203/rs.3.rs-3722826/v1.

E.P.M. conceptualized the project, developed the analytical methods, performed all experiments, analyzed all samples, and wrote the manuscript.

Mueller, E. P., Fenfang Wu, and Alex L. Sessions (2022). "Quantifying Isotopologue Reaction Networks (QIRN): A modelling tool for predicting stable isotope fractionations in complex networks". In: *Chemical Geology* 610. DOI: 10.1016/j.chemgeo.2022.121098.

E.P.M. conceptualized the project, developed the computational software, ran all model simulations and wrote the manuscript.

Mueller, Elliott P., Alex L. Sessions, et al. (2022). "Simultaneous, High-Precision Measurements of $\delta\,^2$H and $\delta\,^{13}$C in Nanomole Quantities of Acetate Using Electrospray Ionization-Quadrupole-Orbitrap Mass Spectrometry". In: *Anal. Chem.* 94. 1092 – 1100.2. DOI: 10.1021/acs.analchem.1c04141.

E.P.M. conceptualized the project, performed analytical experiments and microbial culturing experimental, analyzed the data, and wrote the manuscript.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

C h a p t e r   1

INTRODUCTION

Life, microbial life

Environmental changes occur on timescales that far exceed — by orders of magnitude — our own lifespans. How do we study processes that represent thousands, millions or billions of years of accumulated change when we only glean them for a short few decades? This question is one that Earth and environmental scientists have championed. Lacking physical observations, geoscientists developed chemical techniques that transport us through time. Those same analytical tools also help us capture global processes through the lens of the smallest molecules. An Earth scientist traverses physical and temporal scales with ease. This has always been the most fascinating component of the geosciences to me. That realization was only made more thrilling when I found out that the dominant process controlling these long-term changes was life — microbial life. All over the world, microscopic organisms work together to effect change on a physical scale — a global scale — that far exceeds their own. Since the late 20th century, geochemical studies have come to a consensus on a biological phenomenon: We live in a microbial world.

Microbial life is inextricably linked to the Earth's planetary health. They have witnessed and even caused some of the most dramatic climatic changes in Earth's >4 billion year history. Microbes control our planet's climate by recycling elements like carbon, sulfur, and phosphorus between the continents, oceans, and atmosphere. Yet, they are, by definition, impossible to observe with the naked eye and they mostly grow far slower than we are patient for in the lab. Thus, the vast majority of microbial species have never been observed. As geobiologists and geochemists, we use our expertise to study these enigmatic organisms that operate on physical and temporal scales completely disparate from our own. This thesis represents a small contribution to that grander effort.

Organic acids, the carbon currency of anoxic environments

In the following chapters, I seek to understand the role of microbial metabolisms in organic carbon degradation and its regulation of Earth's past and present climate. Specifically, I try to elucidate the mechanisms of organic degradation in environments that lack oxygen ('anoxic'). In those locations, organic carbon from primary

produced biomass is buried and remains there for decades or centuries. Breakdown of organic material is slow but can release large amounts of greenhouse gases (CO_2 and methane), making it important for short-term climate trends. Burial of primary produced organic matter also controls the net sequestration of carbon dioxide from the atmosphere on geologic timescales, exerting a strong influence on long-term climate trends. I am particularly interested in the initial stages of carbon degradation catalyzed by fermentation, a microbial metabolism that has long been invisible to our myriad techniques. Here, I try to unveil fermentation in the environment by ignoring the cells performing this metabolism and instead focusing on their waste products, organic acids.

Often, we study processes like carbon degradation by analyzing their starting materials (e.g., solid organics) and products (e.g., CO_2 and methane). In my mind, these are the molecular 'banks' of the carbon cycle. The amount of money in each of these banks controls the concentration of greenhouse gases in the atmosphere and thus the Earth's climate. When we look more closely, organic matter cannot transfer money instantaneously to the CO_2 and methane bank. Instead it must pass through an intermediary. Organic matter is first converted to small organic acids by fermenting microbes. Those organic acids are then consumed to generate CO_2 and methane by anaerobic respiration or methanogenesis. Like a dollar bill slipped between banks in a handshake, organic acids are the currency of the carbon cycle. This handshake is happening in diverse environments from the seafloor in the Atlantic to the continental subsurface of North America to the rice paddy fields of Asia. It has likely been happening for billions of years, silently recycling carbon through the biosphere and geosphere.

To understand anoxic carbon cycling, we must resist the urge to conflate abundance with importance. The production and consumption of organic acids are tightly coupled, so their standing concentration rarely accumulates. However, the flux of carbon through this small pool of molecules is huge. In this way, organic acids represent the bottleneck of the carbon cycle (Figure 1.1). Without fermentation, most organic carbon is inaccessible to anoxic microbial communities. In terrestrial wetlands, fermentation products like organic acids are either remineralized to carbon dioxide by anaerobic respiration or converted to methane by methanogenesis. The competition for organic acids controls wetland methane emissions. In marine sediments, the production of organic acids is thought to be the rate-limiting stage of carbon degradation and may influence the global turnover of organic carbon on

geologic timescales.

Given its global importance, fermentation is understudied in biogeochemistry. Fermenting microorganisms are difficult to grow and, to my knowledge, there are no quantitative environmental biomarkers (genomic, lipidomic, etc.) of fermentation. Consequently little is known about how it operates in nature, what limits its activity, or how quickly it turns over carbon. These gaps in our understanding and their implications for Earth's climate regulation motivate my work. In this thesis, I use the isotope ratios of organic acids to understand which direction this molecular currency is being transferred and at what rate (Chapters II, III, IV, and VII). I then apply this understanding to two quite different environments, the Precambrian continental crust (Chapter V) and the gut microbiome of ruminant animals (Chapter VI).
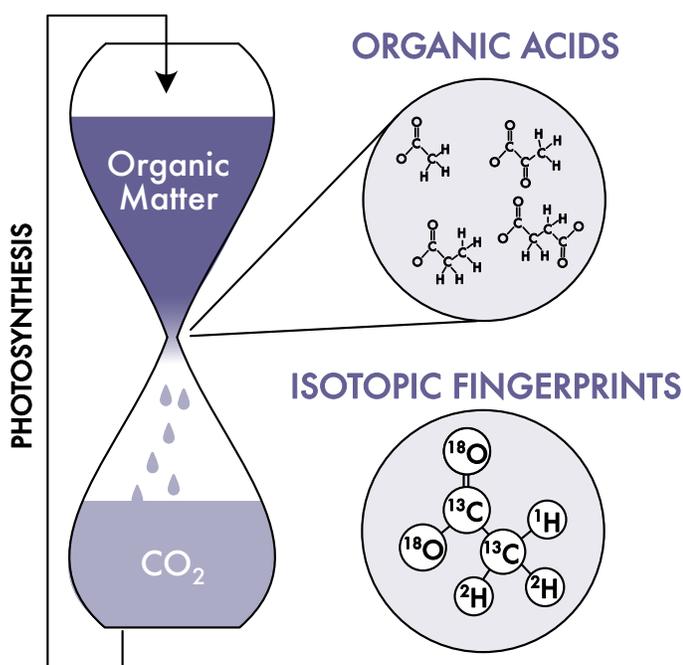
Figure 1.1: The hourglass carbon cycle with organic acids and other small molecules acting as the bottleneck. While low in total abundance, they are rapidly turned over.

Learning to read isotopic stories

Isotope chemistry has been an invaluable tool for studying biogeochemical processes that operate on long timescales. Its use is predicated on small variations in the isotope ratios (i.e., $^{13}C/^{12}C$) of molecules, minerals and bulk materials across the biosphere. For example, during photosynthesis, CO_2 is converted into solid organics with a slight (2%) preference for the $^{12}CO_2$ isotopologue over its heavier, $^{13}CO_2$

counterpart. This causes organic matter to be depleted in ^{13}C relative to the CO_2 substrate. The same idea can be applied to individual reactions, where the molecule serving as substrate is ^{13}C-enriched while the product is ^{13}C-depleted. These isotope ratios can be converted to a more digestible number, known as an isotope composition or a 'delta value'. The $\delta^{13}C$ value of a molecule or material is simply the permil change of its isotope ratio from that of an international reference material. These are expressed in 'permil (‰)' values. For example, the $\delta^{13}C$ value of organic matter is typically 20-30‰ lower than that of CO_2 due to the isotope effect of photosynthesis. Isotope compositions can equally be applied to elements beyond carbon, like hydrogen (δ^2H), nitrogen ($\delta^{15}N$) and sulfur ($\delta^{34}S$).

The isotope composition of a molecule is an inherited property that tells a story about that molecule's origins. Since every organic molecule contains at least two elements with stable isotopes, each molecule has multiple isotopic stories to tell, each of which could lend unique information about that molecule's history. Layered onto this is information about the $\delta^{13}C$ or δ^2H composition of specific atomic sites within a molecule (i.e., site-specific isotopic enrichments). In conventional isotope ratio mass spectrometry (IRMS), molecules are first chemically converted into smaller, gaseous compounds like CO_2 that are easier to manipulate within the instrument. These compounds are then ionized and their isotope ratios are measured. Since all carbon in the original molecule becomes CO_2 , the $\delta^{13}C$ of that CO_2 product represents the analyte's molecular-average $\delta^{13}C$ value. This practice is robust and has delivered decades of useful isotopic measurements, but it eliminates the dense isotopic information stored within molecules. In other words, all of those isotopic stories are condensed into a single statement.

When I arrived at Caltech, new advancements in terms of isotope ratio mass spectrometry were underway to circumvent this issue. My colleagues in John Eiler and Alex Sessions' groups were using Orbitrap-MS to measure the isotopic properties of molecules for the first time. In this technique, molecules are introduced to the mass spectrometer as intact ions. From there, the different isotopologues of those molecules are separated using the Orbitrap's high mass resolution. This meant that they could capture the $\delta^{13}C$ value of the molecule without converting it to CO_2 first. By breaking it into molecular fragments within the instrument and measuring the $\delta^{13}C$ of each fragment, the site-specific isotope composition of those molecules could also be measured. Most of the initial findings on measurements of organic molecules with the Orbitrap were done with gas chromatography (GC) Orbitraps

(Chimiak et al., 2021; Eiler et al., 2017; Hofmann et al., 2020). These are well designed for small sample sizes but biomolecules like amino acids had to be chemically derivatized (e.g., methylation) before they could be analyzed. From there, three things happened at once:

1.) I was getting my first results on the carbon isotope compositions of organic acids and lipids from fermenting bacteria (now Chapter IV) and it was clear to me that carbon isotopes alone might not be enough to elucidate fermentation from other metabolisms (e.g., acetogenesis). I was beginning to consider their hydrogen isotope composition as an extra isotopic story to tell.
2.) John had just purchased an electrospray (ESI) Orbitrap that he wanted to use to measure the isotopic structures of biomolecules without derivatization. No one had made a rigorous effort to measure hydrogen isotopes on molecules yet.
3.) Folks at Caltech and elsewhere were beginning to grapple with the implications of having so much isotopic information about individual molecules. Computational models would need to keep pace with analytical advancements to make sense of this new data.

It was clear from these three events that investing some time into developing an Orbitrap method for measuring the isotopic composition of organic acids could be fruitful. At the same time, I wanted to stay grounded in my questions about microbial carbon degradation, which meant I had to assimilate that new isotopic information into our understanding of fermentative metabolism. These two motivations became the two parts of my thesis.

Part 1 | Acetate — Chapters II, V & VII
Acetate plays a central role in biogeochemistry. It is quantitatively the most important fermentation product and intermediate for carbon cycling in anoxic environments, yet its sources and sinks are hard to distinguish even using its $\delta^{13}C$ value. Acetate is found almost ubiquitously in the environment, but is typically <100 μM in concentration due to rapid consumption by microbial metabolisms. Acetate is also a simple C_2 molecule with hydrogen atoms only on its methyl group. For these reasons, I decided that acetate would be the best analytical target for the first δ^2H measurements on the Orbitrap. In **Chapter II**, I demonstrate that the Orbitrap can indeed make this measurement, using 1000-times less sample than conventional techniques. I further showed that the metabolisms I wanted to distinguish (fermen-

tation and acetogenesis) produced acetate with over 200‰ different $\delta^2 H$ values (Mueller, Sessions, et al., 2022). Alex was determined to measure something in nature to prove the techniques feasibility. It was great pressure to have, because it was no easy feat. Extracting acetate from samples (as was required from ESI-Orbitrap analyses) took a year of method development, but with the help of others in the Sessions and Eiler labs, it worked (**Chapter VII**).

At this point, I was exploring where to apply this technique. Barbara Sherwood Lollar and Verena Heuer had just published a fascinating paper on the carbon isotope composition of acetate in Kidd Creek mine. Kidd Creek is situated in a large rock formation known as the Canadian Precambrian Shield (Sherwood Lollar et al., 2021). Trapped in fractures of the crystalline basement kilometers below the surface are hypersaline fluids that have average residence times of over 1 billion years. These are some of the most ancient groundwaters on Earth. Victoria Orphan offered to introduce me to Barbara, and after six months of tests on synthetic solutions and 'dirty' samples of Kidd Creek waters, Barbara sent me pure samples from three different sites in the mine. Before measuring those samples, I went to visit Kai-Uwe Hinrichs' group at the University of Bremen where Verena works for three months. I wanted to learn about the traditional method for measuring the $\delta^{13}C$ of organic acids, known as liquid chromatography IRMS. Unfortunately, that instrument was out of commission. Kai and Verena helped me work on a separate project where we tested the exchangeability of hydrogen atoms between water and acetate's methyl-group. Those experiments, along with the results we later got from the Kidd Creek samples, turned into a larger collaboration that gave rise to **Chapter V** of this thesis: Turns out, acetate's hydrogen isotope composition can tell us not only about its origins but also its turnover rate in the environment. We applied these ideas to the Precambrian Shield and discovered an active carbon cycle within these ancient waters. Acetate in subsurface fracture fluids was being produced and consumed on extremely long timescales, an organic acid 'cycle' that operates too slowly for us to perceive with concentration measurements alone, even over a ten-year time series (Mueller, Panehal, et al., 2024). The tools I had spent years developing were starting to answer the questions that had initially sparked my interest in geobiology. It was an exciting experience for me.

Part 2 | Fermentation — Chapters III, IV & VI

Stable isotopes are best known as forensic tools, made famous by their use in sports doping investigations (Brooker et al., 2014). In geobiology, they are often used to

a similar effect, distinguishing sources and sinks of metabolites in nature. This has proven to be quite helpful, especially in the rock record, where other proxies (e.g., proteomics, genomics) disappear. Chapters II, V, and VII of my thesis are similar examples of such applications. However, I was eager to push for more quantitative insights from stable isotopes analyses. I drew inspiration from experts like Alex S., John E., John Hayes, Itay Halevy, Ann Pearson, and Kate Freeman, who had all spent parts of their careers trying to disentangle the signals within the stable isotopes composition of minerals and molecules. In an effort to assimilate all of their studies under a single framework, I (along with Alex and Fenfang Wu) developed a modeling software called Quantifying Isotopologue Reaction Networks or QIRN ("churn"). QIRN has some intricacies that differentiate it from previous modeling approaches, but the most important contribution (from my perspective) was its ability to take in and simulate highly complex reaction networks with ease. Fenfang, in particular, was imperative for the initial binary formulation of isotopologues that lend QIRN its expediency and flexibility. QIRN became **Chapter III** of my thesis (Mueller, Wu, and Sessions, 2022). Its development was mirrored by the simultaneous adaptation of Orbitrap technology for isotopic applications, which was quickly generating new data to use as constraints on those models.

The first real application of QIRN came in **Chapter IV**, during which I continued my collaboration with Kai and Verena to measure the carbon isotope fractionation of fermentation. We discovered that fermentation can express strong isotope fractionations, leaving the organic acids either ^{13}C-depleted or ^{13}C-enriched, depending on the type of fermentation pathway it was using. Using QIRN, we isolated the enzymes within those pathways that caused the observed fractionations. Rather than making binary statements about which metabolism generated the organic acid (à la Chapter II), we were observing shifts in enzymatic fluxes within a single metabolism that manifested predictable isotopic fractionations. Additionally, it opened up a new avenue for us to interrogate fermentation, an otherwise invisible metabolism, in modern anoxic environments and in the rock record.

In **Chapter VI**, I focused both the Orbitrap and QIRN onto a single environmental application where fermentation was indisputably the dominant microbial metabolism: Cow rumen. Enteric fermentation is important for understanding climate change (~25% of global methane emissions), but it also represented a relatively simple system where organic carbon was being degraded anaerobically and we could observe it closely using fistulated cows. With this in mind, Alex S. and I reached

out to Matthias Hess at University California (UC) Davis and spun up the final collaboration of my thesis. We used the Orbitrap to measure the $\delta^{13}C$ and δ^2H of organic acids produced by fermentation in the rumen. We then used QIRN and our results from our pure culture experiments in Chapter IV to understand the observed fractionations between cow feed and the excreted organic acids. The ultimate goal is to harness the isotope composition of organic acids as a proxy for the various fermentation pathways that operate in the rumen. These tools are unraveling mechanisms of fermentation and organic degradation that I hope to one day apply in other environments.

Final thoughts

This introduction was meant to give the reader a sense of how seemingly disparate projects came together to speak to a broader problem, one about microbial timescales, molecular currencies, and invisible metabolisms. Splitting the thesis into these two themes helps me do that but the order of my chapters does not fit neatly into parts outlined above. I hope this narrative gives the reader a sense that scientific exploration is neither a linear nor a predictable path. Rarely did projects pane out the way I expected. Research is a balancing act between the questions we are trying to answer and where the data take us. It is also not a path walked alone. Perhaps the strongest thread tying this thesis together is collaboration. Only with the perspectives of my amazing colleagues in Bremen, Davis, Pasadena, and Toronto were these projects possible.

On a final note, I hope this thesis reminds the reader that while some problems are solved by looking at the most abundant molecules, others can only be addressed by interrogating the molecular minority. Like tapping into a phone call between Earth's biosphere and geosphere, carbon currencies give us a window into a broader conversation, one that controls the planet's past, present, and future climates. This thesis represents my translation of but one line from that conversation, and it was a privilege.

References

Brooker, Lance et al. (Oct. 2014). "Stable carbon isotope ratio profiling of illicit testosterone preparations – domestic and international seizures". en. In: *Drug Testing and Analysis* 6.10, pp. 996–1001. ISSN: 1942-7603, 1942-7611. DOI: 10.1002/dta.1533. URL: https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/dta.1533 (visited on 04/19/2024).

Chimiak, L. et al. (Jan. 2021). "Carbon isotope evidence for the substrates and mechanisms of prebiotic synthesis in the early solar system". en. In: *Geochimica et Cosmochimica Acta* 292, pp. 188–202. DOI: 10.1016/j.gca.2020.09.026.

Eiler, John et al. (Nov. 2017). "Analysis of molecular isotopic structures at high precision and accuracy by Orbitrap mass spectrometry". en. In: *International Journal of Mass Spectrometry* 422, pp. 126–142. DOI: 10.1016/j.ijms.2017.10.002.

Hofmann, Amy E. et al. (Nov. 2020). "Using Orbitrap mass spectrometry to assess the isotopic compositions of individual compounds in mixtures". en. In: *International Journal of Mass Spectrometry* 457, p. 116410. DOI: 10.1016/j.ijms.2020.116410.

Mueller, Elliott P., Juliann Panehal, et al. (2024). "Isotopic evidence of acetate turnover in the Precambrian continental crust". In: *Research Square Preprint*. DOI: 10.21203/rs.3.rs-3722826/v1.

Mueller, Elliott P., Alex L. Sessions, et al. (Jan. 2022). "Simultaneous, High-Precision Measurements of $\delta\,^2$H and $\delta\,^{13}$C in Nanomole Quantities of Acetate Using Electrospray Ionization-Quadrupole-Orbitrap Mass Spectrometry". en. In: *Anal. Chem.* 94.2, pp. 1092–1100. DOI: 10.1021/acs.analchem.1c04141.

Mueller, Elliott P., Fenfang Wu, and Alex L. Sessions (Nov. 2022). "Quantifying Isotopologue Reaction Networks (QIRN): A modelling tool for predicting stable isotope fractionations in complex networks". en. In: *Chemical Geology* 610, p. 121098. DOI: 10.1016/j.chemgeo.2022.121098.

Sherwood Lollar, B. et al. (Feb. 2021). "A window into the abiotic carbon cycle – Acetate and formate in fracture waters in 2.7 billion year-old host rocks of the Canadian Shield". en. In: *Geochimica et Cosmochimica Acta* 294, pp. 295–314. DOI: 10.1016/j.gca.2020.11.026.

Chapter 2

HIGH-PRECISION MEASUREMENTS OF ACETATE $\delta^{13}C$ AND δ^2H VIA ELECTROSPRAY ORBITRAP MASS SPECTROMETRY

Elliott P. Mueller[1][*] Alex L. Sessions[1] Peter E. Sauer[2] Gabriella M. Weiss[1,3] John M. Eiler[1]

[1]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA 91125

[2]Department of Earth and Atmospheric Sciences, Indiana University, Bloomington, IN, USA 47405

[3]Astrobiology Center for Isotopologue Research, Department of Geosciences, Pennsylvania State University, State College, PA, USA 16802

*Corresponding Author: emueller@caltech.edu
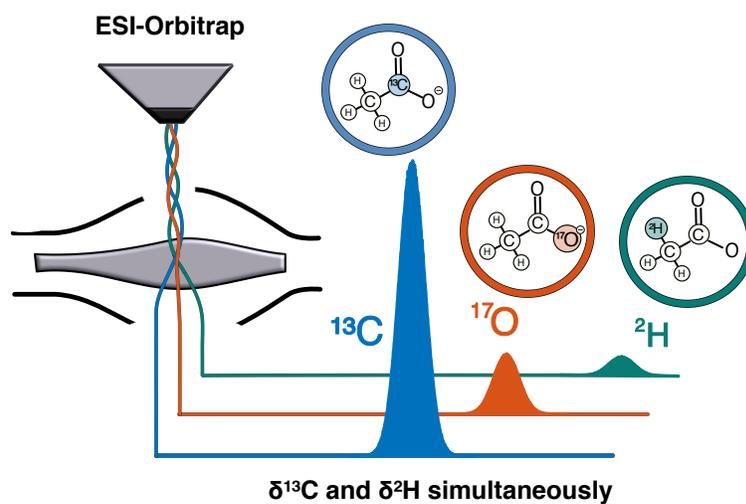
Graphical Abstract of Chapter 2

2.1 Abstract

Stable hydrogen isotope compositions ($^2H/^1$H ratios) have been an invaluable tool for studying biogeochemical processes in nature, but the diversity of molecular targets amenable to such analysis is limited. Here, we demonstrate a new technique for measuring δ^2H of biomolecules via Orbitrap mass spectrometry (MS) using acetate as a model analyte. Acetate was chosen as a target molecule because its production and consumption are central to microbial carbon cycling, yet the mechanisms behind acetate turnover remain poorly understood. δ^2H of acetate could provide a useful constraint on these processes; however, it remains uncharacterized in nature due to analytical challenges. Electrospray ionization (ESI) Orbitrap MS circumvents these challenges and delivers methyl-specific H-isotope compositions of acetate with nanomole sensitivity, enough to enable analyses of environmental samples. This approach quantifies the methyl-specific δ^2H and molecular-average $\delta^{13}C$ of acetate simultaneously while achieving <3‰ and <0.5‰ uncertainty, respectively. Using optimized ionization and Orbitrap parameters, this level of precision is obtained within 15 minutes using only 15 nanomoles of acetate. As a demonstration of our analytical approach, we cultured three acetogenic bacteria and found a large ^{2}H-fractionation between acetate and water (>310‰ depletion) associated with the Wood-Ljungdahl Pathway while fermentation expressed a muted (~80‰) fractionation. With its high precision and sensitivity, Orbitrap MS is a promising tool for investigating these signals in nature after offline purification. Furthermore, the ESI-Orbitrap method presented here could be applied to other molecules amenable to ESI, including central metabolites and sugars, greatly expanding the molecular targets used in hydrogen isotope biogeochemistry.

2.2 Introduction

Hydrogen isotope biogeochemistry draws on a range of molecular targets for studying biogeochemical cycles, paleoclimate, ecology, and other disciplines. However, the structural diversity of these targets is currently limited to those molecules that are abundant and analytically tractable, mainly lipids and hydrocarbons. As such, only a small selection of water-soluble polar metabolites has been surveyed for their hydrogen isotope compositions (Fogel, Griffin, and Newsome, 2016; Smittenberg and Sachs, 2007) The advent of Orbitrap-based isotope ratio measurements with electrospray ionization (ESI) sources could expand the diversity of molecules used for hydrogen isotope biogeochemistry. To test this idea, we selected acetate as a target analyte to develop the simultaneous measurement of $\delta^{13}C$ and δ^2H using an

ESI Orbitrap mass spectrometer (MS).

Acetate was chosen for its significant role in biogeochemical cycles, which could be further elucidated through the lens of its stable hydrogen isotope composition. As an intermediary molecule, acetate acts as both a substrate and product of metabolic processes and is cycled rapidly within microbial communities. The microbial metabolisms that drive acetate cycling are widespread on Earth: In anaerobic environments, acetate is a product of terrestrial plant biomass degradation and a universal substrate for respiring microorganisms. In terrestrial wetlands, acetate fermentation is the largest source of biogenic methane (Conrad, 1999). And in thawing permafrost soils, acetate and other organic acids are thought to constitute a significant portion of the vast carbon reservoirs held beneath the surface, which could become a major source of greenhouse gases (Drake et al., 2015). Due to its common use as a growth substrate, acetate is rapidly consumed and thus held at very low concentrations in nature. This steady-state pool of acetate is continually consumed and replenished by microbial metabolisms, resulting in large fluxes of carbon and low (micromolar) standing concentrations. Its low concentration and ubiquitous utilization have made the flux of acetate through microbial ecosystems difficult to constrain with metabolomics and genomics. However, the natural abundance of stable isotopes could be a useful tool to quantify acetate cycling.

For decades, naturally-occurring stable isotopes have been used to trace and quantify biogeochemical cycles (Hayes, 2001). For example, the carbon stable isotope ($^{13}C/^{12}C$) composition of acetate has been called on to constrain sources of methane and mechanisms of organic remineralization in anoxic sediments (Heuer, Pohlman, et al., 2009; Heuer, Krüger, et al., 2010). Hydrogen isotopes ($^2H/^1H$) can be an even more useful tool for elucidating metabolic processes, because they are highly sensitive to biological fractionations (Zhang et al., 2014; Luxem, Leavitt, and Zhang, 2020; Hattori et al., 2010). Despite its potential utility for distinguishing mechanisms of acetate production and consumption, the hydrogen isotope composition of acetate (and virtually all other water-soluble metabolites) in the environment remains entirely uncharacterized due to analytical challenges.

The most common technique for measuring the hydrogen isotope ratios ($^2H/^1H$) of organic molecules is separation by gas chromatography (GC) followed by pyrolysis and isotope ratio mass spectrometry (IRMS) of the produced H_2 (GC-py-IRMS). This technique is routinely used for volatile and semi-volatile compounds. Deriving isotopic information from polar biomolecules is more difficult using GC-py-IRMS,

because they are not as amenable to gas phase analysis. To measure the hydrogen isotope composition of acetate with GC-IRMS, Hattori et al. (2010) trapped it on a solid phase resin, which was inserted into a GC inlet (solid phase microextraction (SPME)-GC-IRMS). Unfortunately, solid phase extraction required micromoles of acetate, precluding analysis of most environmental samples, and measured the protonated form of acetate (Hattori et al., 2010). Since the carboxylic acid H-atom exchanges with surrounding solvents, samples can only be compared if they have been equilibrated in isotopically identical water.

A more straightforward technique for measuring the isotopic composition of polar biomolecules is to analyze them directly in liquid solutions. Liquid chromatography (LC)-IRMS uses this strategy for acetate and other organic acids, but it cannot measure hydrogen isotopes (Godin and McCullagh, 2011). Nuclear magnetic resonance (NMR) methods can measure hydrogen isotope ratios of the methyl group of acetate solutions without interference from the exchangeable proton on the carboxyl group, but these methods require too much analyte (micromoles) to be useful for environmental applications (Remaud et al., 1992). Here, we present a method using ESI-Orbitrap MS that measures the carbon and hydrogen isotope composition of acetate in solution while circumventing the aforementioned analytical challenges.

This work builds on previous innovations of isotope analyses by Orbitrap MS. In contrast to IRMS methods that chemically convert organic molecules to CO_2 (for $^{13}C/^{12}C$ analysis) or H_2 (for $^{2}H/^{1}H$ analysis), Orbitrap MS isotope ratios are calculated directly from the relative abundances of the isotopologues of a compound's molecular or fragment ions. These ions are produced by ionizing an intact molecule of interest and can be used to deduce molecular average or site-specific isotope compositions (Eiler et al., 2017). Previous Orbitrap MS studies have consistently reported uncertainties on isotope ratios that decreased predictably as more ions were counted (Neubauer, Sweredoski, et al., 2018).Precision was only limited by experimental duration and ion current, both of which have limits proportional to sample size. Deuterium (^{2}H) constitutes less than 0.015% of hydrogen atoms at natural abundance and is the least abundant of the commonly measured stable isotopes. This presents a problem for hydrogen isotope analyses via Orbitrap MS that has yet to be resolved: The method must count enough ^{2}H-isotopologues to achieve low uncertainty on isotope ratios while minimizing experimental duration, and thus sample size, so that it is useful for environmental applications.

The purpose of this report is to explore and document the performance character-

istics of hydrogen isotope analyses of dissolved acetate using Orbitrap MS. Our approach measures the molecular-average carbon and methyl-specific hydrogen isotope composition of acetate by ESI coupled to tandem quadrupole-Orbitrap MS. It optimizes ion filtering and mass analysis parameters to increase the number of isotopologue ions counted per second, resulting in a rapid analysis that delivers hydrogen and carbon isotope compositions that can be accurately calibrated to common isotopic scales defined by Vienna Standard Mean Ocean Water (VSMOW) and Vienna Pee Dee Belemnite (VPDB) reference materials, respectively. Results from this study suggest that natural abundance $\delta^2 H$ and $\delta^{13} C$ values can be simultaneously quantified by Orbitrap MS in 15 minutes with only 15 nanomoles of acetate while achieving precisions better than 0.5‰ and 3‰, respectively.

2.3   Methods

Acetate Standards

Sodium acetate (>99%) was obtained from three commercial sources (Fisher Scientific; Mallinckrodt Chemicals; Allied Chemical) to create standards with a range of carbon and hydrogen isotope compositions (Table S1). Their synthetic or biological origins are unknown. Stock solutions (2M) were made up in deionized water (Milli-Q IQ 7000 Purification). The stocks were brought to pH 8-9 with 5M sodium hydroxide (Mallinckrodt Chemicals) solution to avoid evaporation of protonated acetic acid. Working solutions were created by diluting the stocks to 50 μM with HPLC grade methanol (VWR Life Science). These working solutions were directly infused into the instrument.

Instrumentation and Data Analysis

A Q Exactive HF Orbitrap mass spectrometer (Thermo Fisher Scientific) was coupled to a Fusion 101 (Chemyx, Inc.) syringe pump equipped with 500 μL syringe (Hamilton Robotics). The pump continuously infused sample into an ESI source with an HESI-II probe (Thermo Fisher Scientific) at 20 μL/min. Ionization parameters for the HESI probe were as follows: polarity: negative, spray voltage: -3.0 kV; auxiliary gas flowrate: 3; capillary temperature: 320°C; sheath gas flow rate: 10; sweep gas flow rate: 0; spray current: Õ.2 mA; S-Lens RF level: 50; auxiliary gas heater temperature: 100°C.

The following mass spectrometer parameters were used throughout this study: scan type: selected ion monitoring (SIM); scan range: 57-62 m/z; microscans: 1; maximum injection time: 1000 ms; automatic gain control (AGC): 10^6; mass

resolution: 60,000 (unless otherwise stated). This mass resolution refers to the full width at half-maximum mass-resolving power for an ion of 200 m/z; the actual mass resolving power changes by a factor of based on the mass-to-charge ratio (m/z) of a given ion. The resolution of the monoisotopic acetate ion was between 110,500 and 441,900 depending as the resolution setting was adjusted from 60,000 to 240,000.

FTStatistic software (an unpublished, proprietary program created by Thermo Fisher Scientific) was used to retrieve ion signal intensities and noise from the 'RAW' files that are automatically generated by the instrument control software. Ion intensities are not a direct measurement of the number of ions entering the analyzer (a quantity we refer to as 'counts'); rather, the Orbitrap observes all ionic species simultaneously and quantifies their signals (image current observations) through a Fourier Transform (FT) deconvolution algorithm that integrates a period of observations with its length set by the resolution. To convert ion intensities to ion counts (Equation 1), the signal-to-noise ratio (S/N_P) of a given spectral peak in a single scan was integrated over the length of the scan defined by the resolution of the measurement (R, 60,000 for most experiments reported here) and a reference resolution (R_N, 120,000). Finally, the signal was converted to counts (N) by applying an empirical factor (C_N) derived by Makarov and Denisov (2009) and the charge of the ion (z). (Makarov and Denisov, 2009; Eiler et al., 2017).

$$N = \left(\frac{S}{N_P} \right) \left(\frac{C_N}{z} \right) \sqrt{\frac{R_N}{R}} \qquad (2.1)$$

Isotope Working Standards

Batches of working 2M standard stock solutions in Milli-Q water (Table S1) at pH 8-9 were flash frozen in liquid nitrogen. To ensure homogeneity, stocks were kept as frozen aqueous solutions. Aliquots from these stocks were taken by thawing them at room temperature, inverting the vials to homogenize, aliquoting and immediately re-freezing the stocks. Each standard was measured via Elementary Analyzer (EA) coupled to an Isotope Ratio Mass Spectrometer to independently determine their isotopic compositions relative to the VSMOW and VPDB reference materials. Carbon isotope compositions were measured with combustion EA-IRMS. Hydrogen isotope compositions were measured with thermal conversion EA-IRMS using an elemental chromium catalyst as discussed by Gehre et al., 2015. See Supplementary Information for details. These standards are referred to as AcSt, AcA, and AcB.

Method Overview

A 50 μM solution of an acetate working standard in methanol was infused continuously into the Orbitrap. Carbon and hydrogen isotope ratios of the working standard were measured every ~175 ms over a five-minute acquisition. The same analysis was then performed on a purified acetate sample of interest, and this sample/standard comparison was repeated a total of three times (totaling 15 minutes per sample). After each acquisition (of sample or standard), the syringe and PEEK tubing that connects the syringe to the ESI source were flushed with 1.5 mL of methanol before being refilled with the new analyte solution. Blank acquisitions consistently reported < 5% of 59 m/z signal from a 50μM sample, and no correction to account for this blank signal was performed.

Calculation of Isotope Ratios

Uncorrected carbon and hydrogen isotope ratios (R_{OT}) were calculated directly from isotopologue counts (N) obtained by Orbitrap analyses:

$$^{2}R_{OT} = \frac{(^{12}C_2^2H^1H_2^{16}O_2)_N}{(^{12}C_2^1H_3^{16}O_2)_N} \tag{2.2}$$

$$^{13}R_{OT} = \frac{(^{12}C^{13}C^1H_3^{16}O_2)_N}{(^{12}C_2^1H_3^{16}O_2)_N} \tag{2.3}$$

Samples were compared to laboratory working standards to derive delta values:

$$\delta = \frac{R_{OT,sample}}{R_{OT,standard}} - 1 \tag{2.4}$$

where $R_{OT,standard}$ is an average of the two working standard measurements analyzed before and after each sample replicate. This bracketing strategy, as in many other IRMS measurements, was implemented to correct for subtle instrument drift. Delta values can also be calculated relative to the international reference materials VSMOW and VPDB. To do so, a standard with known isotope compositions from EA-IRMS analyses is measured on the Orbitrap to calculate a correction factor (C_F) for the instrumental fractionation:

$$C_F = \frac{R_{OT,standard}}{R_{IRMS,standard}} \tag{2.5}$$

Delta values in the VSMOW and VPDB scales are then calculated as:

$$\delta = \frac{R_{OT,sample}/C_F}{R_{ref}} \tag{2.6}$$

where R_{ref} is 0.01118 and 0.00015576 for carbon and hydrogen isotope systems, respectively (Brand et al., 2014). Delta values presented here are relative to VPDB or VSMOW reference frames unless otherwise stated.

The uncertainty associated with internal reproducibility of isotope ratios throughout a single acquisition, referred to here as the acquisition error (AE), is calculated as the standard error of isotope ratios across scans:

$$\sigma_{AE} = \frac{\sqrt{\sum_{i=1}^{n}(R_{OT,i} - \overline{R_{OT}})}}{n} \tag{2.7}$$

where i is the scan number; $R_{OT,i}$ is the isotope ratio for scan i; is the averaged isotope ratio across all scans of an acquisition; and n is the total number of scans in an acquisition. The total uncertainty of a single measurement is the experimental reproducibility and is defined as the reproducibility of isotope ratios across multiple acquisitions of the same sample. Here, it is first calculated as a standard deviation in comparisons with acquisition error and later calculated as standard error for defining the uncertainties on sample measurements. Experimental reproducibility can equivalently be expressed in terms of sample-standard comparisons in delta notation.

Acetate Isolation Protocol

To measure the isotopic composition of acetate from biological cultures or environmental samples, acetate was separated from matrices with a combination of clean-up steps and ion chromatography. For details on culturing conditions, see Supplementary Information. Ag/Ba/H 2.5cc cartridges (Thermo Scientific Dionex™) were rinsed with 500 mL of deionized water prior to sample introduction. Blank measurements following this rinse did not detect acetate above detection limits. Sample solution (0.5 mL) was pressed through an Ag/Ba/H 2.5cc cartridge at 1mL/min to remove chloride and sulfate from solution without binding acetate. An additional 1mL of deionized water was collected after samples eluted to remove any residual acetate. Samples were buffered with Tris buffer (ThermoFisher) and sodium hydroxide to a pH of 7-8 and then evaporated to 0.8 mL under N_2. Samples were then injected onto a Dionex™ Integrion HPIC Ion Chromatograph equipped with an

AS11-HC column (2 mm x 250 mm), an 822 μL sample loop, and a Dionex™ 2mm ASRS suppressor. Sample injection volumes ranged from 300-500 μL with a push partial injection method. A KOH eluant gradient from 1mM to 60mM was used to separate acetate from other ions in solution. Acetate was collected via manual fraction collection into combusted glass vials, evaporated under N_2, and redissolved in HPLC-grade methanol. Samples were then analyzed on the ESI-Orbitrap as described above.

2.4   Results and Discussion

Acetate Ionization

With an ESI source, polar molecules can be introduced as liquid solutions without preparatory chemical derivatization or conversion. In negative polarity mode, electrospray ionization produces intact acetate ions by deprotonating the carboxylic acid moiety (Figure 2.1A). As a result, the singly-^{13}C-substituted isotopologue captures acetate's molecular average carbon isotope composition while the singly-[2]H-substituted isotopologue measures the hydrogen isotope composition of acetate's methyl site. Methyl-bound hydrogen atoms on free acetate in solution are not readily exchangeable with water and thus preserve isotopic information relevant for geobiological applications (Sessions et al., 2004).

To precisely quantify the isotopologues of acetate and calculate isotope ratios, we first optimized the ESI spray conditions. An infusion flow rate of 20 μL/min was necessary to stabilize the total ion current (TIC) below 3% (1 SD, scan-to-scan), which constitutes a significant increase from previous related methods (Neubauer, Crémière, et al., 2020; Neubauer, Sweredoski, et al., 2018). When flow rate was decreased to 10 μL/min, TIC fluctuations increased to 20%. We hypothesize that the relative TIC instability is caused by the poor ionization yield of acetate.

In blanks and samples, we observed small amounts of dissolved, singly charged carbonate ions (59.98 Da). Oxyanions are a common feature of background spectra in negative ESI mode (Keller et al., 2008). While carbonate ions are readily distinguished by the Orbitrap, their mass is too close to that of the [13]C- and [2]H-isotopologues of acetate (~60.02 Da) for them to be separated by the quadrupole MS (a mass filter). It has been previously shown that the presence of background ions in the Orbitrap can alter measured isotope ratios even when they are mass-resolved from the ions of interest (Hofmann et al., 2020). Methanol, which has a low solubility for carbonate species, was therefore used as a solvent to mitigate this problem.

Adding water to the solvent mixture caused the carbonate ion abundance to increase significantly. Lowering pH to convert carbonate ions to carbonic acid attenuated the signal of acetate's molecular ion, potentially because of its protonated state. While the carbonate ion is still present when using methanol as the solvent, its relative abundance is consistent and is well below the threshold that impacts isotope ratios defined by previous work (Hofmann et al., 2020).
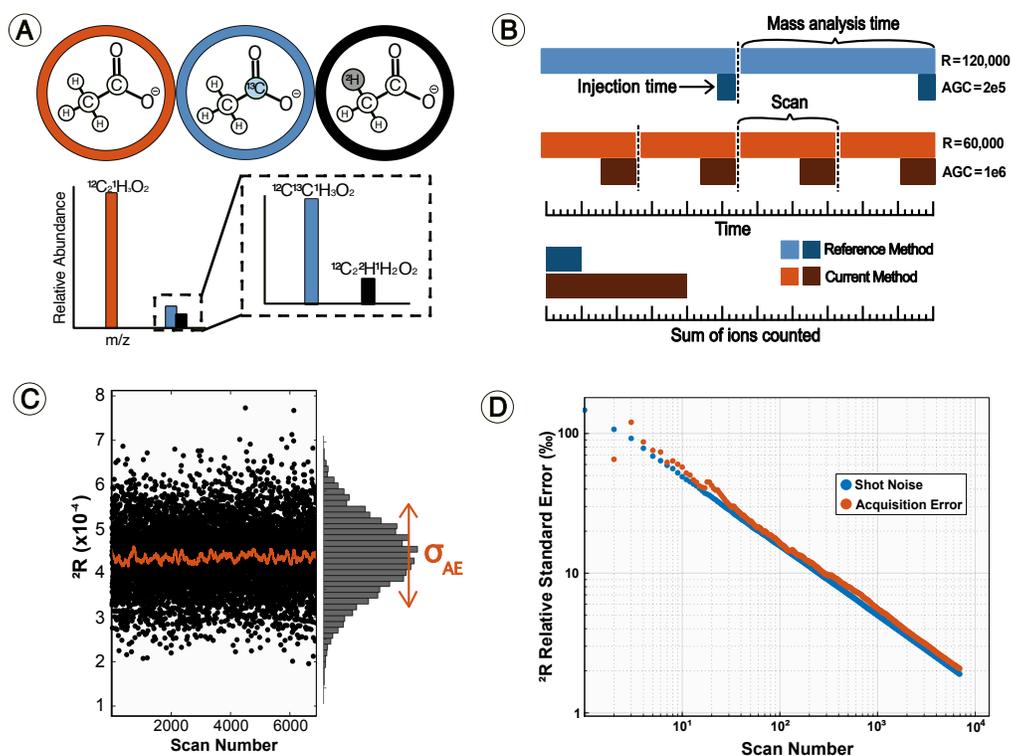
Figure 2.1: (A) Simultaneous measurement of all three isotopologues of acetate during a single acquisition. (B) Diagram of Orbitrap scan cycles, which include filling the C-trap with ions (injection time), injecting into the Orbitrap (dotted lines), and mass analysis. Our method (red) optimizes ion counting efficiency by increasing both the rate of Orbitrap scans (with a lower resolution) and injection time (with higher AGC targets). A reference method (blue) at the 120,000 resolution setting and 2×10^5 AGC target — representative settings for related prior Orbitrap methods — is presented for comparison. (C) Monitoring the isotope ratios from each scan produces a normal distribution of $^2R_{OT}$ with an AE. A moving average of 100 scans is shown as the red line. (D) Throughout an acquisition, the AE is consistently proportional to the limit set by shot noise (i.e., counting statistics).

Maximizing Ion Counts

The precision of isotope ratios measured by Orbitrap MS commonly approaches the limit set by counting statistics (shot noise limit) (Eiler et al., 2017; Neubauer, Crémière, et al., 2020). In other words, the uncertainty of an isotope ratio measurement is inversely proportional to the square root of the number of ions that are counted. Longer analysis times therefore achieve greater precision but also consume more sample. This is potentially problematic for quantifying exceptionally rare isotopologues (< 0.015% relative abundance), which will require much larger samples to count the same number of rare isotopologues. To minimize counting time, we optimized the Orbitrap's ion counting efficiency (the ratio of ions counted in the Orbitrap to total ions emitted from the ESI source) by adjusting quadrupole filtering, C-trap collection, and Orbitrap mass analysis settings. In the following sections, we explain each of these modifications and their effects on the precision and accuracy of the method. With these optimized MS parameters, we performed rapid quantification of ^{2}H- and ^{13}C-isotopologues simultaneously using only nanomoles of acetate.

AQS Range As ions exit the ESI source, they are filtered through an advanced quadrupole system (AQS) to select the mass range that will enter the Orbitrap. We set the AQS range between 57 and 62 Da to measure all relevant isotopologues of acetate simultaneously while minimizing transmission of background ions. Previous experiments have used an even narrower (1 Da) range that isolated only M+1 ions.[19] The goal of that approach was to increase the counting rate of singly-substituted isotopologues by excluding the more abundant, monoisotopic ions. Here we found that counting ^{2}H-isotopologue ions is more efficient using a 5 Da window. Narrowing the AQS window further resulted in longer injection times that ultimately limited scan cycle rates and negated any benefit to ion counting efficiency. This was another consequence of acetate's low ionization yield in the ESI source. Although including the monoisotopic ions with a wider AQS window does indeed decrease the relative abundance of deuterated isotopologues entering the Orbitrap, space-charge effects (interactions between proximal ions that attenuate their signal) on ^{2}H-isotopologues are diminished as a result. The benefits of narrowing the AQS were therefore outweighed by the negative consequences of longer injection times and space-charge effects.

AGC Target Control After AQS filtering, ions accumulate in an ion trap ("C-trap") up to a threshold set by the automated gain control (AGC) target. Trapped ions are

then injected into the Orbitrap and analyzed. The C-trap remains idle until mass analysis is nearly complete at which point it samples from the ion current again. Since mass analysis time typically takes 10-100x longer than filling the C-trap (injection time), the C-trap is idle for much of the run and only a small fraction of the ion current is analyzed, resulting in a low ion counting efficiency (Figure 2.1B).

To maximize counting efficiency, we increased the AGC target to 10^6, pushing beyond the limit suggested by prior isotopic studies (Eiler et al., 2017). By expanding the injection time, more of the ion current is sampled per scan (Figure 2.1B). Previous studies have avoided using high AGC targets because it increases the density of ions in the Orbitrap, which could result in a preferential loss of rare isotopic species: As ions oscillate within the Orbitrap, they proceed along spiraling trajectories that resemble rings. Large rings of abundant ions can influence the trajectories of less abundant ions. The smaller ion rings begin to converge on the same frequency as the larger rings in a process known as coalescence. Coalescence is thought to occur between ions with similar masses and may become more prevalent as ion density increases (Eiler et al., 2017). This is observed as a shift in the reported mass of the rare ion towards that of the more abundant ion. It also decreases the number of rare ions counted per scan (Khodjaniyazova et al., 2018). Despite the increase in AGC target, the mass difference between the ^{2}H- and ^{13}C-isotopologues of acetate does not deviate from the theoretical mass difference by more than two ppm. We also observed a commensurate increase in the ^{2}H-isotopologue counts with a higher AGC target.

We hypothesize that coalescence at high AGC targets is mitigated by including the more abundant monoisotopic ion. Coalescence on the ^{2}H ion ring could theoretically worsen as the relative abundance of the ^{13}C ion ring increases, because these two isotopologues are < 0.01 Da different by mass and the ^{13}C ion ring is significantly more abundant. With a wide AQS window, the monoisotopic ion becomes the most abundant species in the Orbitrap and the ^{13}C-isotopologue is a minor peak (< 3%). By decreasing the relative abundance of ^{13}C-isotopologue ions in the Orbitrap, the ^{2}H-isotopologue experiences less severe coalescence and AGC target can be increased without attenuating the ^{2}H-signal.

Mass Resolution Orbitrap mass analysis time is directly proportional to the desired mass resolution (M/ΔM) and is the rate-limiting step in scan cycles (Figure 2.1B). Thus, the most effective way to increase ion counting efficiency is to decrease resolution to the minimum needed to resolve nominally isobaric isotopologues.
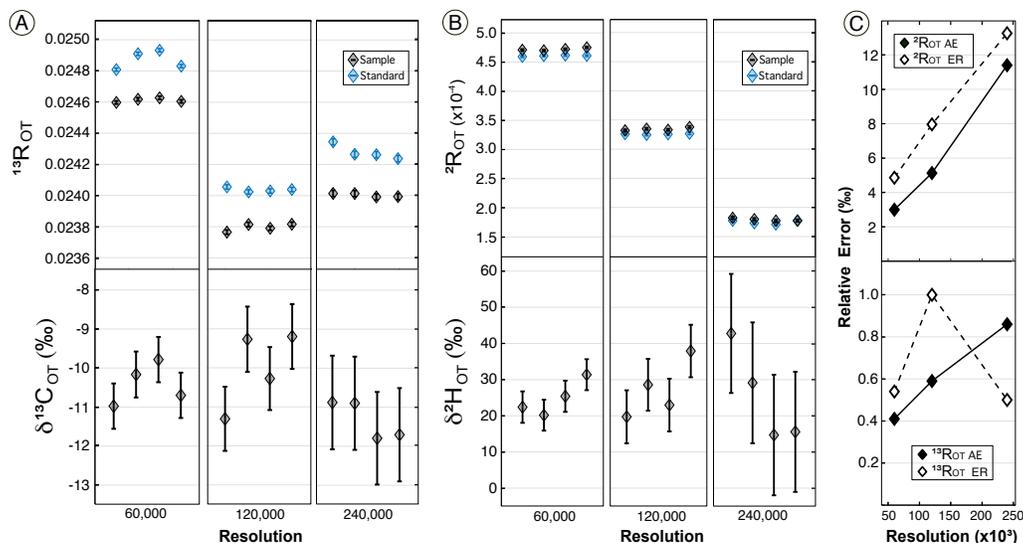
Figure 2.2: Replicate acquisitions reproduced isotope ratios within AE. Despite significant changes in measured isotope ratios of (A) carbon and (B) hydrogen with changing mass resolution, sample-standard comparisons reproduced similar δ^2H $_{OT}$ and $\delta^{13}C$ $_{OT}$ values at every resolution tested. Delta values are relative to a laboratory working standard, not VPDB or VSMOW. (C) At each resolution, the ER of replicate acquisitions was calculated as standard deviation and the AE of a single acquisition was calculated as standard error. Each is calculated for $^2R_{OT}$ (top) and $^{13}R_{OT}$ (bottom).

Since acetate has a relatively small mass, ^2H- and ^{13}C-isotopologues were resolved at a relatively low resolution setting of 60,000, providing a significant increase in scan cycle rate as compared to most previous Orbitrap isotope ratio measurements (Figure 2.1B).

Higher resolutions will be needed to quantify the ^2H-isotopologues of larger molecules, exposing their ions to longer mass analysis times on each scan. Since rare ions have faster orbital decay rates than abundant ions, higher mass resolution could in theory suppress the abundance of rare isotopologues and alter isotope ratios (Khodjaniya-zova et al., 2018; Gordon and Muddiman, 2001). To test this, we measured acetate standards with a range of resolution settings while holding all other Orbitrap MS parameters constant (Figure 2). These experiments confirmed that measured isotope ratios can change dramatically with resolution. This was most evident in the hydrogen isotope measurements, where the $^2R_{OT}$ value at 60,000 resolution was nearly equivalent to the expected value at natural abundance (natural abundance ~4.7e-4). As resolution increased, measured isotope ratios decreased three-fold for both sample and standard (Figure 2.2B). These changes were well outside ex-

perimental uncertainty of replicates and are consistent with our inference that the ^{2}H-isotopologue ion ring decays faster in the Orbitrap than that of the more abundant monoisotopic ion.

Despite such instrument artifacts, sample-standard comparisons reported the same $\delta^{13}C_{OT}$ and δ^2H_{OT} values across all resolution settings (Figure 2.2A,B). Thus, we found that orbital decay effects are reproducible enough to be readily corrected by sample/standard comparison with identical analysis parameters. This suggests that future measurements of rare isotopologues in larger molecules can potentially use the same Orbitrap MS method presented here and simply increase the resolution setting to the desired level.

Measurement Precision

With its modified AQS, AGC, and resolution settings, this Orbitrap method optimizes the counting efficiency of ^{2}H-isotopologues. To test whether these enhancements improve the internal precision of the measurement, we monitored $^2R_{OT}$ over a 20-minute direct infusion. Throughout the acquisition, each scan cycle corresponded to a discrete measurement of the sample's isotope composition. These individual $^2R_{OT}$ values produced a normal distribution around an average with a standard error referred to here as the 'acquisition error' (Figure 2.1C). As scans accumulated, more isotopologue counts were recorded and acquisition error decreased. Over the course of the infusion, acquisition error of $^2R_{OT}$ was nearly equivalent to the theoretical limit imposed by counting statistics (Figure 2.1D). We have yet to observe a deviation from this trend, even in hour-long acquisitions that consumed 60 nanomoles of acetate. In experiments where acquisition time was held constant (7 minutes) while resolution sequentially increased, the acquisition error rose proportionally with resolution due to the lower ion counting efficiency (Figure 2.2C). However, in each experiment, the acquisition error remained nearly equivalent to its respective shot noise limit, which also increased with resolution. Thus, the uncertainty on isotope ratios measured by a single acquisition using the instrument and methods presented here is only limited by sample size. We chose a default acquisition time of five minutes for future experiments as it achieves an acquisition error comparable to other methods of measuring hydrogen isotope ratios ($\sim$3‰, 1SE) and uses only five nanomoles of analyte.

Characterizing the acquisition error does not capture the full experimental uncertainty. Errors associated with memory effects, reloading samples between acqui-

sitions, and sample introduction (i.e. syringe pump pressure) may add additional inaccuracy to the measurement. A more useful metric for total error is the standard deviation of R_{OT} across analytical replicates, referred to here as the 'experimental reproducibility'. To quantify the experimental reproducibility of $^{13}R_{OT}$ and $^{2}R_{OT}$ measurements, we alternated between an acetate sample and standard (AcSt, see Methods), measuring four replicates of each at three resolutions (Figure 2.2). For $^{2}R_{OT}$ values, the experimental reproducibility was consistently offset from the acquisition error by $\sim 2\permil$, even as the latter increased with resolution. The experimental reproducibility on carbon isotope ratios did not have a similarly clear relationship with acquisition error. However, at a resolution of 60,000, the default setting used to measure acetate's isotope composition, experimental reproducibility was nearly equivalent to acquisition error (Figure 2.2C).

These data suggest that replicate analysis does not introduce significant error despite many opportunities for instrument drift or other analytical artifacts to appear. With the current procedure, syringes filled with acetate are reloaded by hand and manually placed in the syringe pump. The instrument is put in standby mode between infusions, and the same syringe is used for sample and standard. However, we did not observe significant memory effects or monotonic drift in R_{OT} values in these experiments. Furthermore, sample-standard comparisons shown in $\delta^{13}C$ $_{OT}$ and $\delta^{2}H$ $_{OT}$ were reproducible without correcting for instrument drift (Figure 2.2). More robust technologies with dual inlet syringe pumps for sample and standard would avoid human error in sample loading, sample carry over between infusions, and idle instrument time. Such modifications could lower the experimental reproducibility of the measurements even further (Hilkert et al., 2021). Nevertheless, our current measurements achieves internal and external uncertainty (1 SE) in $\delta^{2}H$ values of less than $3.0\permil$, equivalent or lower than current NMR and SPME-GC-IRMS methods while requiring 50-1000x less sample (Hattori et al., 2010; Remaud et al., 1992).

Measurement Accuracy

Isotope ratios measured with mass spectrometers can diverge from their true values due to instrumental fractionation (a.k.a. bias). To quantify and correct for such effects in the Orbitrap, the isotopic composition of an acetate working standard (AcSt) was characterized by EA-IRMS against international reference materials. Two additional acetate samples with distinct isotopic compositions (AcA and AcB) were measured with both EA-IRMS and Orbitrap MS. During Orbitrap experiments, AcSt was run alongside AcA and AcB to calculate their $\delta^{13}C$ and $\delta^{2}H$ values on
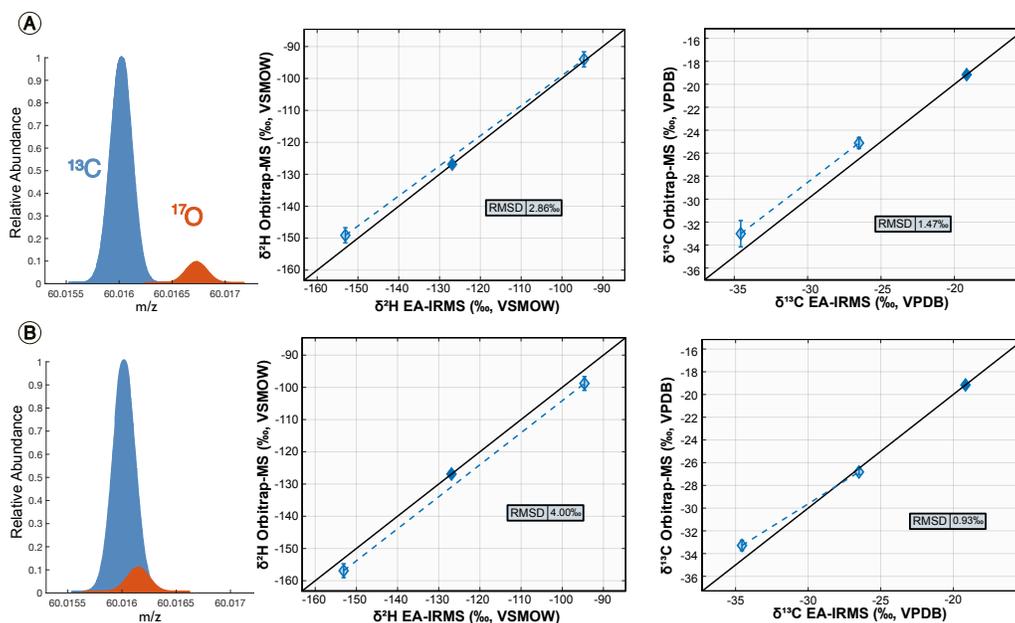
Figure 2.3: δ^2H and $\delta^{13}C$ values of acetate samples (AcA and AcB, shaded diamonds) were measured by both EA-IRMS and Orbitrap MS to estimate the accuracy of the latter. An acetate working standard (AcSt, solid diamond) was used to correct for instrumental fractionation during Orbitrap MS experiments. The accuracy of the method was sufficient to measure natural isotopic fractionations at both 120,000 (A) and 60,000 (B) resolution settings, despite the ^{13}C- and ^{17}O peaks merging at the lower resolution. RMSDs between EA-IRMS and Orbitrap MS measurements are shown on the plots. Error bars represent the ER (1 SE) of triplicate analyses.

the VDPB and VSMOW scales (Equations 5 and 6). Orbitrap acquisitions were 10 minutes long and each sample was measured in triplicate. We take the root-mean-square-deviation (RMSD) between Orbitrap MS and EA-IRMS measurements as an estimate of the accuracy of the Orbitrap method (Figure 2.3). At 120,000 resolution (Figure 2.3A), Orbitrap MS measurements of δ^2C $_{\text{VSMOW}}$ and $\delta^{13}C$ $_{\text{VPDB}}$ yielded 3.0‰ and 1.5‰ RMSD, respectively. This level of accuracy is more than sufficient to elucidate natural isotopic fractionations. However, scale nonlinearities are likely to exist and will decrease accuracy as standard and sample become farther apart in isotope ratio.

Accuracy calibrations against EA-IRMS measurements were first performed at 120,000 resolution to resolve all singly-substituted acetate isotopologues, including the ^{17}O-isotopologue. However, at 60,000 resolution, the ^{17}O and ^{13}C-isotopologues are not resolved and appear as a single mass spectral peak, potentially affecting the accuracy of $^{13}R_{\text{OT}}$ measurements. If the $\delta^{17}O$ compositions of the sample and stan-

dard are significantly different, the effect of merging the ^{13}C and ^{17}O-peaks would increase the ^{13}R$_{OT}$ values of the sample and standard by different magnitudes and alter the measured $\delta^{13}C$ $_{VPDB}$ value. However, when the accuracy calibration was repeated at 60,000 resolution, the $\delta^{13}C$ $_{VPDB}$ values of AcA and AcB did not change significantly and the RMSD (1.0‰) decreased (Figure 2.3B).

These results suggest that differences between the oxygen isotope compositions of our AcA, AcB and AcSt samples were not significant. The oxygen atoms on acetate may have equilibrated with the water used as a solvent in stock solutions. If true, the three samples would have nearly identical oxygen isotope compositions and no change with resolution would be observed. Given that the oxygen atoms are located on a carboxylic acid, exchange with water is possible (Redington, 1976). We suggest that methods developed in the future perform similar accuracy tests that resolve the ^{17}O isotopologue. Ignoring this species lowers the minimum resolution needed and dramatically improves precision for a given fixed analytical duration.

Acetate Isolation

Next, we examined the influence of matrix effects on ionization and designed a workflow for purifying acetate via ion chromatography. When acetate standards were analyzed in methanol solutions containing 1 mM of different salts, the molecular ion intensity decreased by ten-fold (Figure 2.6). This attenuation may be the result of matrix effects in the ESI source where acetate ions could cluster with cations. Regardless of the cause, most natural samples of acetate will have to be purified from their matrices before infusion into the Orbitrap MS. We thus designed an ion chromatography preparatory method to remove interfering ionic species from natural samples and to concentrate pure acetate.

We first tested whether IC separation induces isotopic fractionation by spiking water with 100 μM of the acetate standard, AcB. It was then isolated from the solution in triplicate and analyzed on an ESI-Orbitrap. After sample clean-up and IC collection, AcB was depleted by 6.1‰ (± 2.4‰) and 3.9‰ (± 1.4‰) in δ^2H and $\delta^{13}C$, respectively, compared to the known values. Hydrogen isotope fractionation is marginal given the uncertainty on replicates. However, the carbon isotope fractionation is significant. When 50 μM solutions of AcB in water were dried down, redissolved in methanol, and measured on the Orbitrap, $\delta^{13}C$ values depleted by only 0.5‰ (± 0.9‰, SD triplicate experiments), suggesting that evaporation is not the source of carbon isotopic fractionation. The goal of future studies will be to identify the cause

of this fractionation and systematically correct for it using isotopically characterized standards prepared alongside samples.
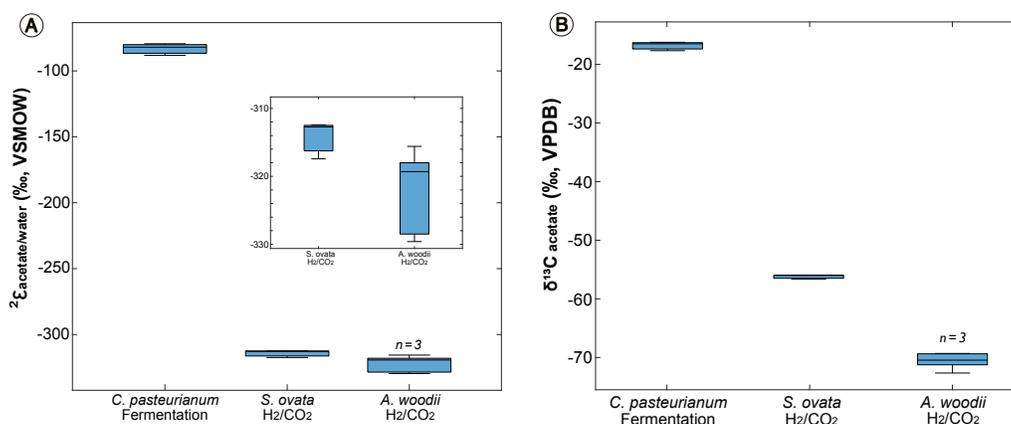
Figure 2.4: Hydrogen isotope fractionation between media water and acetate (A) and carbon isotope composition of acetate (B) from two autotrophic acetogens (*A. woodii* and *S. ovata*) and a fermenting bacterium (C. pasteurianum), where n is the number of biological replicates. The inset is an enlarged plot of the fractionation from the two autotrophic acetogens.

Biological Fractionations of Acetogenesis

As an applied demonstration of our new analytical approach, we measured the acetate produced by cultures of three acetogenic bacteria that utilize two different metabolic pathways, fermentation and autotrophic acetogenesis. These metabolisms comprise the primary biological sources of acetate in anaerobic environments.[7] Our results indicate a significant difference in $\delta^2 H$ of acetate they produced (Figure 2.4). *Clostridium pasteurianum* was grown with glucose as the sole carbon source in a defined medium under fermenting conditions, while *Acetobacterium woodii* and *Sporomusa ovata* produced acetate from H_2 and CO_2 through the Wood-Ljungdahl Pathway (WLP) (Supplementary Tables 2.2-2.3). The two autotrophic acetogens were grown on different sources of CO_2 that had not been isotopically characterized; however, they both excreted highly ^{13}C-depleted acetate characteristic of the WLP (Figure 2.4B) (Blaser, Dreisbach, and Ralf Conrad, 2013).

All three cultures expressed a measurable fractionation of hydrogen isotopes between the water they were grown in (-94‰) and the excreted acetate (Figure 2.4A). Notably, both autotrophic acetogens exhibited a large -320‰ fractionation, whereas the fermenter, *C. pasteurianum*, expressed a much smaller -80‰ fractionation. To confirm these results, *A. woodii* was grown in triplicate; biological reproducibility

of this isotopic signal was within analytical precision (< 3‰). This finding is also consistent with the data of Valentine et al., 2004, who observed that the lipids of *Sporomusa* sp – another autotrophic acetogen – were highly ^{2}H-depleted (-410‰) relative to growth water. We hypothesize that a large hydrogen isotope fractionation associated with the acetogenic WLP may be a robust biosignature of that metabolism. If correct, then the large (>200‰) offset in δ^2H values between WLP and fermenting acetogens could be exploited to understand their relative importance in natural environments after thorough investigation of sink processes that may also fractionate the hydrogen isotope composition of acetate. Orbitrap MS not only has the precision to probe this range, but also the sensitivity to investigate these signals in a myriad of natural environments.

2.5   Conclusions

ESI-Orbitrap MS is an attractive technique for measuring the carbon and hydrogen isotope ratios of small polar organic molecules, including acetate. Here we measured $\delta^{13}C$ and δ^2H of acetate in 15 minutes with an accuracy of 1‰ and 4‰, respectively, using only 15 nanomoles of analyte for triplicate analyses. The Orbitrap-based method does require offline purification of acetate, which we accomplish via ion chromatography. With this workflow, we identified a large biological signal between fermenting and autotrophic acetogenic bacteria, two major sources of acetate in anaerobic environments. Given the excellent sensitivity of the Orbitrap, application of such measurements to natural samples should be straightforward. These isotopic signals could be used to elucidate the role that acetate production and consumption plays in ocean sediments, terrestrial wetlands, or other important settings. The potential for advancements in TIC stabilization with lower infusion rates and acetate isolation from complex matrices warrants future work as these improvements would lower sample requirements even further.

Our analytical approach is likely to be useful for other molecules amenable to electrospray ionization (i.e. lipids, metabolites, organic acids, etc.). As such, ESI-Orbitrap MS greatly expands the diversity of molecular targets that can be used to study hydrogen isotope biogeochemistry. The internal mass spectrometric parameters that achieved rapid, precise, and accurate results in this study can also be applied to other rare (<0.015%) isotopic species, such as clumped isotopologues or the minor isotopes of oxygen and sulfur.

Acknowledgements

2.6   Supplementary Materials

Carbon isotope compositions of three acetate working standards were measured on a combustion Elemental Analyzer (EA) IsolinkTM system coupled to a Delta V Plus IRMS with a ConFlo IV Universal interface (all Thermo Scientific, Bremen, Germany). Carbon isotope ratios of combusted CO_2 were calibrated to the VPDB scale with internal CO_2 reference gas (-12.04‰) measurements before and after sample peaks. External standards valine (USGS74, 9.3‰), glycine (USGS64, -40.81‰), sucrose (NIST, -10.47‰), and caffeine (IAEA, -27.77‰) were used to correct for further instrumental fractionations associated with combustion.

Hydrogen isotope compositions were measured on a temperature conversion (TC) EA IRMS. Samples were weighed into silver capsules and dried in a glass vacuum desiccator for hours. The desiccator was flushed with nitrogen gas before being opened. Silver capsules were quickly crushed and sealed in a Zero-Blank autosampler, which was flushed with helium. The instrument was set to a furnace temperature of 1100°C, a GC oven temperature of 40°C, and a flow rate of 70 mL/min. Samples were calibrated to the VSMOW scale by analyzing USGS77 (polyethylene powder) and C36 n-alkane #2 provided by Arndt Schimmelman (Indiana University).

Description of working standards

Table 2.1: Acetate working standards measured by EA-IRMS

Standard	$\delta^{13}C$ (‰, VPDB)	δ^2H (‰, VSMOW)	Source
AcSt (Sodium acetate)	-19.2 ± 0.1‰	-127 ± 2‰	Malinckdrodt
AcA (Sodium acetate)	-26.5 ± 0.1‰	-153 ± 2‰	Fisher Scientific
AcB (Sodium acetate)	-34.6 ± 0.1‰	-95 ± 1‰	Allied Chemicals
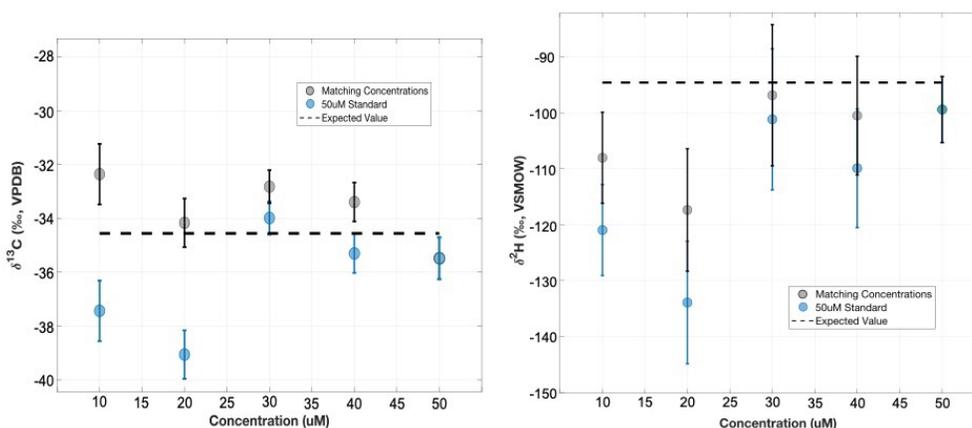
Figure 2.5: An acetate standard (AcB) was measured against working standard AcSt across a range of concentrations. AcSt was either kept at a constant 50 μM concentration (blue) or was diluted to match the concentration of AcB (black). Accuracy decreased when the solutions were < 30 μM, potentially due to the larger influence of background acetate at low concentrations

Table 2.2: Bacterial strains surveyed in this study.

Species	Strain	Source	Source Location
Sporomusa ovata	DSM 2662	Volker Müller	Goethe University, Germany
Acetobacterium woodii	DSM 1030	Volker Müller	Goethe University, Germany
Clostridium pasteurianum	LMG 3285	DSMZ	Brunswick, Germany

Culture Conditions:

Acetobacterium woodii and Sporomusa ovata were both grown in 25 mL balch tubes with 10 mL of media under 20 psi 80:20 H_2:CO_2 at pH 7 and 30 $\delta^{13}C$. Cultures were shaken at 180 rpm and grew to completion over one day. Media aliquots were taken in stationary phase, filtered through a 0.2 μm filter, and frozen.

Clostridium pasteurianum was grown in 1L sealed culture bottles with 100 mL of media under an N2 atmosphere at 37 $\delta^{13}C$. Cultures were not shaken. Media aliquots were taken in end-log phase, filtered through a 0.2 μm filter, and frozen.

Table 2.3: Media composition for bacterial cultures

Component	A. woodii	S. ovata	units
KH_2PO_4	0.2	2.15	g/L
NH_4Cl	0.25	—	g/L
Na_2HPO_4	—	9.16	g/L
NH_4SO_4	—	3	g/L
NaCl	1.16	—	g/L
$MgSO_4 \cdot 7H_2O$	1.45	—	g/L
$CaCl_2$	0.11	0.013	g/L
KCl	0.5	—	g/L
$KHCO_3$	6	—	g/L
$Na_2S \cdot 9H_2O$	0.3	—	g/L
Na_2SeO_3	17	—	μg/L
$Na_2WO_4 \cdot 2H_2O$	33	—	μg/L
$FeSO_4 \cdot H_2O$	2	—	mg/L
$FeCl_3 \cdot 6H_2O$	2027	2027	μg/L
H_3BO_3	30	30	μg/L
$MnCl_2 \cdot 4H_2O$	100	100	μg/L
$CoCl_2 \cdot 6H_2O$	190	190	μg/L
$NiCl_2 \cdot 6H_2O$	24	24	μg/L
$CuCl_2 \cdot 2H_2O$	2	2	μg/L
$ZnCl_2$	68	68	μg/L
Riboflavin	100	100	μg/L
Biotin	30	30	μg/L
Thiamine HCl	100	100	μg/L
L-Ascorbic acid	100	100	μg/L
d-Ca-pantothenate	100	100	μg/L
Folic acid	100	100	μg/L
Nicotinic acid	100	100	μg/L
4-aminbenzoic acid	100	100	μg/L
pyridoxine HCl	100	100	μg/L
Lipoic acid	100	100	μg/L
Thiamine pyrophosphate	100	100	μg/L
Cyanocobalamin	10	10	μg/L
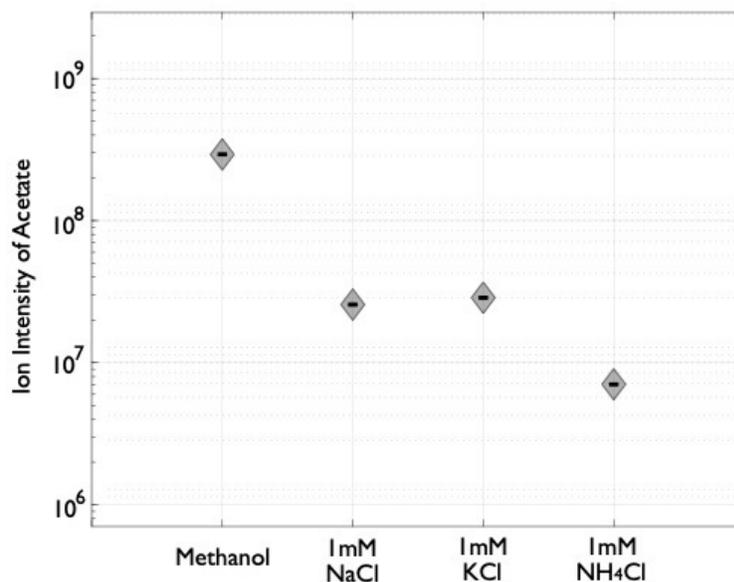
Figure 2.6: Addition of 1 mM of NaCl, KCl and NH$_4$Cl to a 50 μM acetate solution in methanol reduces the ionization of acetate bymore than 10-fold.

References

Blaser, Martin B., Lisa K. Dreisbach, and Ralf Conrad (Mar. 2013). "Carbon Isotope Fractionation of 11 Acetogenic Strains Grown on H $_2$ and CO $_2$". en. In: *Appl Environ Microbiol* 79.6, pp. 1787–1794. DOI: 10.1128/AEM.03203-12.

Brand, Willi A. et al. (Mar. 2014). "Assessment of international reference materials for isotope-ratio analysis (IUPAC Technical Report)". en. In: *Pure and Applied Chemistry* 86.3, pp. 425–467. DOI: 10.1515/pac-2013-1023.

Conrad, R (Mar. 1999). "Contribution of hydrogen to methane production and control of hydrogen concentrations in methanogenic soils and sediments". en. In: *FEMS Microbiology Ecology* 28.3, pp. 193–202. DOI: 10.1111/j.1574-6941.1999.tb00575.x.

Drake, Travis W. et al. (Nov. 2015). "Ancient low–molecular-weight organic acids in permafrost fuel rapid carbon dioxide production upon thaw". en. In: *Proc. Natl. Acad. Sci. U.S.A.* 112.45, pp. 13946–13951. DOI: 10.1073/pnas.1511705112.

Eiler, John et al. (Nov. 2017). "Analysis of molecular isotopic structures at high precision and accuracy by Orbitrap mass spectrometry". en. In: *International Journal of Mass Spectrometry* 422, pp. 126–142. DOI: 10.1016/j.ijms.2017.10.002.

Fogel, Marilyn L., Patrick L. Griffin, and Seth D. Newsome (Aug. 2016). "Hydrogen isotopes in individual amino acids reflect differentiated pools of hydrogen from food and water in *Escherichia coli*". en. In: *Proc. Natl. Acad. Sci. U.S.A.* 113.32. DOI: 10.1073/pnas.1525703113.

Gehre, Matthias et al. (May 2015). "On-Line Hydrogen-Isotope Measurements of Organic Samples Using Elemental Chromium: An Extension for High Temperature Elemental-Analyzer Techniques". en. In: *Anal. Chem.* 87.10, pp. 5198–5205. DOI: `10.1021/acs.analchem.5b00085`.

Godin, Jean-Philippe and James S. O. McCullagh (Oct. 2011). "Review: Current applications and challenges for liquid chromatography coupled to isotope ratio mass spectrometry (LC/IRMS)". en. In: *Rapid Comm Mass Spectrometry* 25.20, pp. 3019–3028. DOI: `10.1002/rcm.5167`.

Gordon, Eric F. and David C. Muddiman (Feb. 2001). "Impact of ion cloud densities on the measurement of relative ion abundances in Fourier transform ion cyclotron resonance mass spectrometry: experimental observations of coulombically induced cyclotron radius perturbations and ion cloud dephasing rates". en. In: *J. Mass Spectrom.* 36.2, pp. 195–203. DOI: `10.1002/jms.121`.

Hattori, Ryota et al. (June 2010). "Measurement of the Isotope Ratio of Acetic Acid in Vinegar by HS-SPME-GC-TC/C-IRMS". en. In: *J. Agric. Food Chem.* 58.12, pp. 7115–7118. DOI: `10.1021/jf100406y`.

Hayes, John M (2001). "Fractionation of the Isotopes of Carbon and Hydrogen in Biosynthetic Processes". en. In.

Heuer, Verena B., Martin Krüger, et al. (Jan. 2010). "Experimental studies on the stable carbon isotope biogeochemistry of acetate in lake sediments". en. In: *Organic Geochemistry* 41.1, pp. 22–30. DOI: `10.1016/j.orggeochem.2009.07.004`.

Heuer, Verena B., John W. Pohlman, et al. (June 2009). "The stable carbon isotope biogeochemistry of acetate and other dissolved carbon species in deep subseafloor sediments at the northern Cascadia Margin". en. In: *Geochimica et Cosmochimica Acta* 73.11, pp. 3323–3336. DOI: `10.1016/j.gca.2009.03.001`.

Hilkert, Andreas et al. (July 2021). "Exploring the Potential of Electrospray-Orbitrap for Stable Isotope Analysis Using Nitrate as a Model". en. In: *Anal. Chem.* 93.26, pp. 9139–9148. DOI: `10.1021/acs.analchem.1c00944`.

Hofmann, Amy E. et al. (Nov. 2020). "Using Orbitrap mass spectrometry to assess the isotopic compositions of individual compounds in mixtures". en. In: *International Journal of Mass Spectrometry* 457, p. 116410. DOI: `10.1016/j.ijms.2020.116410`.

Keller, Bernd O. et al. (Oct. 2008). "Interferences and contaminants encountered in modern mass spectrometry". en. In: *Analytica Chimica Acta* 627.1, pp. 71–81. DOI: `10.1016/j.aca.2008.04.043`.

Khodjaniyazova, Sitora et al. (Feb. 2018). "Characterization of the Spectral Accuracy of an Orbitrap Mass Analyzer Using Isotope Ratio Mass Spectrometry". en. In: *Anal. Chem.* 90.3, pp. 1897–1906. DOI: `10.1021/acs.analchem.7b03983`.

Luxem, Katja E., William D. Leavitt, and Xinning Zhang (Sept. 2020). "Large Hydrogen Isotope Fractionation Distinguishes Nitrogenase-Derived Methane from Other Methane Sources". en. In: *Appl Environ Microbiol* 86.19. Ed. by Haruyuki Atomi, e00849–20. DOI: 10.1128/AEM.00849-20.

Makarov, Alexander and Eduard Denisov (Aug. 2009). "Dynamics of ions of intact proteins in the Orbitrap mass analyzer". en. In: *J. Am. Soc. Mass Spectrom.* 20.8, pp. 1486–1495. DOI: 10.1016/j.jasms.2009.03.024.

Neubauer, Cajetan, Antoine Crémière, et al. (Feb. 2020). "Stable Isotope Analysis of Intact Oxyanions Using Electrospray Quadrupole-Orbitrap Mass Spectrometry". en. In: *Anal. Chem.* 92.4, pp. 3077–3085. DOI: 10.1021/acs.analchem.9b04486.

Neubauer, Cajetan, Michael J. Sweredoski, et al. (Nov. 2018). "Scanning the isotopic structure of molecules by tandem mass spectrometry". en. In: *International Journal of Mass Spectrometry* 434, pp. 276–286. DOI: 10.1016/j.ijms.2018.08.001.

Redington, Richard L. (Jan. 1976). "Kinetics of oxygen-18 exchange between carboxylic acids and water". en. In: *J. Phys. Chem.* 80.3, pp. 229–235. DOI: 10.1021/j100544a003.

Remaud, Gérald et al. (1992). "A coupled NMR and MS isotopic method for the authentication of natural vinegars". en. In: *Fresenius J Anal Chem* 342.4-5, pp. 457–461. DOI: 10.1007/BF00322207.

Sessions, Alex L. et al. (Apr. 2004). "Isotopic exchange of carbon-bound hydrogen over geologic timescales 1 1Associate editor: J. Horita". en. In: *Geochimica et Cosmochimica Acta* 68.7, pp. 1545–1559. DOI: 10.1016/j.gca.2003.06.004.

Smittenberg, Rienk H. and Julian P. Sachs (Oct. 2007). "Purification of dinosterol for hydrogen isotopic analysis using high-performance liquid chromatography–mass spectrometry". In: *Journal of Chromatography A* 1169.1-2, pp. 70–76.

Valentine, D. L. et al. (July 2004). "Hydrogen isotope fractionation during H_2 /CO_2 acetogenesis: hydrogen utilization efficiency and the origin of lipid-bound hydrogen". en. In: *Geobiology* 2.3, pp. 179–188. DOI: 10.1111/j.1472-4677.2004.00030.x.

Zhang, Xinning et al. (Apr. 2014). "Nitrogen isotope fractionation by alternative nitrogenases and past ocean anoxia". en. In: *Proc. Natl. Acad. Sci. U.S.A.* 111.13, pp. 4782–4787. DOI: 10.1073/pnas.1402976111.

Chapter 3

QUANTIFYING ISOTOPOLOGUE REACTION NETWORKS (QIRN): A GENERALIZABLE FORWARD MODEL FOR PREDICTING ISOTOPE FRACTIONATIONS WITHIN COMPLEX REACTION NETWORKS.

Elliott P. Mueller[1*], Fenfang Wu [1], Alex L. Sessions[1]

[1]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA 91125

Graphical Abstract of Chapter 3

3.1 Abstract

Natural-abundance stable isotope compositions are powerful tools for understanding complex processes across myriad scientific disciplines. However, quantitative interpretation of these signals often requires equally complex models. Previous stable isotope models have treated isotopic compositions as intrinsic properties of molecules or atoms (e.g. $\delta^{13}C$, ^{13}R, etc.). This has proven to be a computationally efficient but inflexible approach. Here, we present a new isotope modelling software tool that combines computational strategies used in metabolic modeling with an understanding of natural isotope fractionations from the geosciences, called Quantifying Isotopologue Reaction Networks (QIRN, "churn"). QIRN treats isotopic properties as distributions of discrete isotopologues, i.e. molecules with different numbers and distributions of isotopic substitutions. This approach is remarkably generalizable and computationally tractable, enabling models of reaction networks with unprecedented complexity. QIRN parameterizes reactions as rate law equations with distinct isotopologues as the reactants and products. Isotope effects are implemented as small changes to the relevant isotopologues' rate constants. Running this model forward in time gives the numerical solution for steady state isotopologue abundances. Different subsets of the isotopologue population can then be sampled to quantify numerous isotopic proprieties simultaneously (i.e. compound-specific, site-specific, and multiply-substituted isotope compositions). Furthermore, QIRN can model any physical, chemical or biological process as reversible or irreversible. As such, it incorporates both kinetic and equilibrium isotope effects. It can be readily applied to any isotope system (i.e. C, N, O, etc.), though at present can only track 2 isotopes of one element at a time. Given its generalizability, QIRN has a diverse range of applications. To demonstrate the flexibility and efficiency of QIRN, we reconstructed previous (intrinsic-property) models of sulfate reduction, abiotic amino acid synthesis, lipid biosynthesis, and photosynthesis. In these examples, QIRN consistently reproduced outputs from prior models and predicted isotopic anomalies that have been measured in nature. With its new approach to isotope modelling, QIRN will expand the potential complexity of modelled reaction networks, help predict isotopic signals that can direct experimental efforts, and provide a more efficient means of modeling emerging isotopic properties such as 'clumped isotopes'.

3.2 Introduction

The study of stable isotopes at natural abundance has found its way into a diverse suite of scientific fields including biogeochemistry, paleoclimate, planetary science, ecology, and many more. Ratios of stable isotopes (i.e. $^{13}C/^{12}C$) are typically used as otherwise inert tracers of physical, chemical and biological processes that can discriminate between heavy and light isotopes. Such reactions can sometimes be detected in nature by measuring the deficit of the heavy (rare) isotope in the product compared to the reactant, e.g. the characteristic ^{13}C depletion of Calvin Cycle products relative to source CO_2. However, these signals are just as often overprinted by additional source and sink reactions, which further fractionate the product of interest. To correctly understand isotopic signals, they must be placed in the context of their broader reaction networks. Unfortunately, with a limited number of isotopic constraints, natural systems must be simplified to make interpretations (e.g., "light" carbon is biogenic). In response to these limitations, analytical capabilities have advanced, unlocking yet more nuanced isotopic properties for analysis (Neubauer et al., 2018; Yoshida and Toyoda, 2000; Wieloch et al., 2018; Douglas et al., 2017). As more detailed analyses are developed and more complicated reaction networks are considered, the interpretation – rather than availability – of isotopic data becomes limiting. Recent publications make clear the role for isotopic models in filling this gap (Gilbert et al., 2012; Gilbert, 2021; Tang et al., 2017; Sim et al., 2019; Jonathon Gropp, 2022).

The current paradigm for quantifying isotopes in reaction networks treats isotopic composition as an intrinsic property of molecules or materials (i.e., ^{13}R or $\delta^{13}C$). Following the framework outlined in previous work (reviewed by John M Hayes, 2001), isotopic compositions are assigned to reactants and products in reaction networks. Reactions themselves are assigned fractionation factors, which transform the isotope ratio between the reactant and product by a constant proportion. This framework is implemented mathematically by finding the analytical solution to a system of equations that describe the flux and fractionation of each reaction in the network, yielding the steady-state isotopic compositions of products. For simple networks, a solution can be obtained by hand. Even for more complicated networks, the approach benefits from very fast computational times, allowing for rapid quantification of large parameter spaces. However, it has several significant drawbacks. First, these equations are relatively inflexible. The addition of new reactions or branching points in the network requires that the equations be rewritten and solved anew. Thus the approach becomes quite cumbersome with increased

complexity and/or variability in the reaction network. Second, for a model to track site-specific isotopic compositions, a separate fractionation factor must be assigned to each atomic site and additional networks must be created if the atomic sites have different sources. This is of particular relevance as position-specific isotope analysis is becoming more analytically accessible (Neubauer et al., 2018; Wieloch et al., 2018; Nimmanwudipong et al., 2016; Thomas, Freeman, and Arthur, 2009). Third, the intrinsic-property framework struggles to clearly describe isotopic properties that involve interactions between more than one atom, such as multiply-substituted isotopologues, particularly when intramolecular isotopic heterogeneities may influence the stochastic reference composition.

A recent adaptation of the intrinsic property approach has been to employ numerical methods in forward models of networks (i.e. Jonathon Gropp, 2022; Wilkes and Pearson, 2019; Goldman et al., 2019). In these studies, simple equations incorporating relative fluxes and isotope effects describe how the isotopic compositions of reactants and products evolve with time. Eventually, the system reaches steady state at which point isotope ratios or delta values are reported. This approach eliminates the need for analytical solutions, making it much easier to change the network topology. However, similar to the analytical approach, forward models must be restructured for every atomic site of interest, and properties involving multiple isotopic substitutions are still problematic. In both prior approaches, the difficulties that come with increased network or molecular complexity (e.g. time to construct models) often force models to parameterize molecules as having average isotopic properties (removing information gained by site-specific isotope measurements) or to neglect processes known to be important in nature.

Major advances in computational tools for reaction network analysis have also been realized in the metabolic modelling community. It has been demonstrated that when ^{13}C-labelled substrates are fed to cells, the ^{13}C-isotopomer patterns of downstream metabolites are uniquely determined by the cell's metabolic fluxes (Wiechert, 2001). After this discovery, mathematical tools were developed to perform the forward calculation from reaction network topology to isotopomer distributions. K. Schmidt et al., 1997 provided a simple matrix-based operation to solve for the steady-state isotopomer patterns of metabolites. Their algorithm takes advantage of a binary notation that allows the program to track isotopomers autonomously. Forward calculations are then iterated to match measured isotopic labelling of central metabolites (Christensen and Nielsen, 1999). Although these

metabolic models do track the presence or absence of an isotope label at individual molecular positions, they do not account for isotopic fractionations induced by reactions (although see Millard, Portais, and Mendes, 2015 and Wasylenko and Stephanopoulos, 2013), which is required to consider natural abundance variations.

In the current work, we have merged the isotopomer tracking algorithms from metabolic modelling with the nuanced understanding of isotope fractionation from the geoscience and ecology communities. The approach is packaged into a flexible tool called Quantifying Isotopologue Reaction Networks (QIRN, pronounced "churn"). QIRN treats substrates as distributions of discrete isotopologues and isotopomers (for simplicity we will hereafter collectively refer to these species as isotopologues). For example, between the seven major stable isotopes of carbon, hydrogen, and oxygen, acetate ($C_2H_4O_2$) has over 1000 isotopologues. All of these exist in nature and interact with biological, chemical and physical processes individually and, potentially, at different rates. QIRN tracks every isotopologue of interest (e.g., the distinct C isotopologues) as separate species. The relative abundances of these isotopologues define the isotopic compositions of a molecule. The discrete isotopologue approach has previously been used to express isotopic compositions (i.e. (Goldman et al., 2019), (Ono, Rhim, and Ryberg, 2022), (Rhim and Ono, 2022)). Notably, Millard, Portais, and Mendes, 2015 built a model of fully-labelled substrates in metabolic pathways that incorporated kinetic isotope effects. QIRN uses similar isotopologue tracking strategies, but focuses on natural abundance isotope compositions and fractionations of interest to the geoscience and ecology communities.

In QIRN simulations, isotopologues undergo reactions individually. These reactions are parameterized as first- or second-order rate equations. Small differences between isotopologues' rate constants control the site-specific and clumped isotope effects of reactions. QIRN runs a forward finite-difference model of the user-defined reaction network to find a steady-state numerical solution for the substrates' isotopic compositions, eliminating the need for analytical solutions. The numerical method alongside binary notation for isotopologues enables QIRN to rapidly generate models of nearly any complexity. These models can adapt to network topology changes even in the event of atomic rearrangements within and between molecules. For a given binary isotope system (C, N, S, etc.), any isotopic property of a molecule can be determined simultaneously by summing different subsets of the isotopologue population, including compound-specific, site-specific and clumped isotope com-

positions. Indeed, users can now interpret numerous isotopic constraints at once without restructuring the network or rerunning the model.

A key question at the onset of QIRN's development was whether the computational expense of explicitly tracking isotopologues would be prohibitive. As we show below, QIRN is indeed slower than intrinsic-property models but can nevertheless simulate networks of modest complexity (>50 reactions) to steady state (10,000 time steps) in <15 seconds using only a laptop computer. To demonstrate the flexibility and efficiency of QIRN, we have reconstructed previous (instrinsic property) models of isotope fractionations in abiotic and metabolic networks and reproduced their predictions. We find that QIRN's explicit-isotopologue approach coupled to its graphic user interface (GUI) significantly expedites network construction and interpretation.

3.3 Methods

Model Overview

QIRN creates finite-difference numerical models of user-defined reaction networks run through a graphic user interface (GUI) with Python scripts and comma-separated-value (csv) input files. Nodes in these networks are substrates, which QIRN represents as pools of isotopologues. The distribution of isotopologues defines a compound's isotopic composition. Isotopologue abundances are relative to the overall concentration of a substrate. Reactions connecting nodes in the network are parameterized as first- or second-order mass action rate laws with separate rate constants for each isotopologue. At each timestep, the fluxes through every reaction are calculated, and the abundances of every isotopologue in the network are updated to reflect these fluxes. Isotope effects are intrinsically included through the different rate constants of each isotopologue.

As a forward finite-difference model, QIRN can be used in two ways: 1.) To solve time-varying concentrations, fluxes and isotope compositions of reaction networks parameterized by mass action rate laws. In these cases, rate constants and substrate concentrations must be known in absolute terms, i.e. with units of sec^{-1} and μM, respectively. The time-step length is user-defined and remains constant over a given model run. However, care is required in selecting a time step that is short enough to produce a stable numerical solution, depending on the varying magnitudes of the rate constants. We note that many other flux models already exist that can accurately model time-varying fluxes and concentrations, and this is not the primary focus of

QIRN.

2.) To solve for isotopic compositions in steady state models. In these cases, only relative rate constants need be known and so can be treated as effectively unitless. Likewise, concentration and time-step have arbitrary units. Models are run out to steady-state concentrations and isotope compositions, but of course the absolute magnitudes of substrate concentrations and fluxes in such a model are meaningless.

Regardless of the application, the user inputs three sets of parameters: 1.) The initial conditions (isotope composition and concentrations) of compounds in the network, 2.) The reaction rate constants for each reaction 3.) The site-specific isotope effects for any fractionating reactions. In this study, we modelled only steady state reaction networks. We adopted a constant concentration and isotopic composition of one or more starting reactants as our boundary condition, allowing products to accumulate and intermediates to reach steady-state. Rate constants were set between 0 and 10 for easier user-handling. These rate constants produced stable solutions with a 0.1 time-step. Once steady state was reached, the isotope compositions (delta values) of the products were calculated from their isotopologue abundances. QIRN does have the ability to implement other boundary conditions and to report time-varying concentrations and isotope compositions, but these operations were not used in the following study.

Data Structure

The distribution of isotopologues for each compound is stored as 1xN arrays, in which each array element stores the abundance of a particular isotopologue, and the index of that array element encodes the identity of the isotopologue. Substrate arrays have a length of 2^n, where n = number of atoms of interest in the molecule. While a variety of strategies for mapping array indices to isotopologue identity are possible, we adopted a particularly efficient one in which the binary representation of the index number describes the number and locations of rare isotope substitutions. Each digit of the binary number corresponds to a particular atom (e.g., C-1 or C-2, etc) and a zero at that digit represents the common isotope (^{12}C) while a one represents the rare isotope (^{13}C). Thus, in a C_3 molecule, array index 0 (binary 000) corresponds to the monoisotopic (all-^{12}C) isotopologue, index four (100) corresponds to a ^{13}C substitution at C-1, index five (101) to ^{13}C substitutions at both C-1 and C-3, etc.

$$
\begin{bmatrix} I_0 \\ I_1 \\ I_2 \\ I_3 \\ \vdots \\ I_8 \end{bmatrix} = \begin{bmatrix} 000 \\ 001 \\ 010 \\ 011 \\ \vdots \\ 111 \end{bmatrix} = \begin{bmatrix} {}^{12}C\,{}^{12}C\,{}^{12}C \\ {}^{12}C\,{}^{12}C\,{}^{13}C \\ {}^{12}C\,{}^{13}C\,{}^{12}C \\ {}^{12}C\,{}^{13}C\,{}^{13}C \\ \vdots \\ {}^{13}C\,{}^{13}C\,{}^{13}C \end{bmatrix} \tag{3.1}
$$

QIRN allows users to input the isotopic compositions of starting reactants in the common delta notation, e.g. $\delta\,{}^{13}C$ values. At the start of a model run, QIRN must first convert these delta values to abundances of individual isotopologues, as follows. For a compound with a non-zero initial abundance, QIRN first assigns all the isotopologue abundances to unity. It then calculates their abundances by looping through the digits of the binary number for each isotopologue (Equation 2). When it encounters a "1", it multiplies the isotopologue abundance by the fractional abundance of the rare isotope at that site (^{13}F, Equation 4). When it encounters "0", it multiplies by that of the abundant isotope. Finally, it multiplies the entire array by the user-defined concentration for that compound, resulting in an array of isotopologue abundances normalized to the assigned total concentration. The arrays of downstream substrates with no initial concentration begin with abundances of zero for every isotopologue and are populated as the model runs forward in time.

$$
I_4 = 1\ 0\ 0 = {}^{13}C\,{}^{12}C\,{}^{12}C = {}^{13}F_{C1} \times {}^{12}F_{C2} \times {}^{12}F_{C3} \times A_{int} \tag{3.2}
$$

where A_{int} is the initial concentration of the compound of interest.

The isotopic fractional abundances for each site (i.e. $^{13}F_{C1}$) of starting substrates are calculated from user-defined site-specific delta values. These delta values are first converted to isotope ratios using a reference isotope ratio (R_{ref}) and then to fractional abundances.

$$
{}^{13}R_{C1} = \left(\frac{\delta^{13}C_{C1}}{1000} + 1\right) \times {}^{13}R_{ref} \tag{3.3}
$$

$$
{}^{13}F_{C1} = \frac{{}^{13}R_{C1}}{{}^{13}R_{C1} + 1} \tag{3.4}
$$

Isotopologoue Flux Calculations

QIRN reactions are parameterized as first- or second-order mass action rate laws. In its current state, the number of reactants or products cannot exceed two each; however, this can be increased in future updates. At each time step, the flux of each reaction is calculated using the abundance of isotopologues from the preceding step. The flux is used to update the relevant abundances for the next timestep. Flux and isotopologue abundance change with time, while the time step and rate constants are unchanging. For given isotopologues of reactants A and B, the flux (J) is calculated as:

$$J_t = k \times [I_A]_t \times [I_B]_t \times dt \tag{3.5}$$

$$[I_A]_{t+1} = [I_A]_t - J_t \tag{3.6}$$

where k is the reaction rate constant and $[I_A]_t$ and $[I_B]_t$ are the abundances of the two isotopologues at timestep t. The timestep (dt) in our models was set to 0.1 and was constant throughout each model run. If a reaction only has one reactant, the term I_B is set to unity.

The product(s) of the reaction (C) are also updated:

$$[I_C]_{t+1} = [I_C]_t + J_t \tag{3.7}$$

Implementing Isotope Fractionation Factors
Kinetic Isotope Effects

With reaction fluxes parameterized as mass action rate laws, kinetic isotope effects can be implemented as small changes in the rate constants for each isotopologue. When a molecule undergoes a reaction, the array of isotopologue abundances representing that molecule is multiplied by an array of rate constants. Both of these arrays are 2^n long where n is the number of of atoms (carbon, nitrogen, sulfur, etc.) of interest in the molecule and $N = 2^n$ is the resulting number of isotopologues that must be tracked. They are arranged in a specific order based on the binary numbering system described above.

$$
\begin{bmatrix} I_0 \\ I_1 \\ \vdots \\ I_{N-1} \end{bmatrix}_{flux} = \begin{bmatrix} I_0 \\ I_1 \\ \vdots \\ I_{N-1} \end{bmatrix}_{conc} \times \begin{bmatrix} k_0 \\ k_1 \\ \vdots \\ k_{N-1} \end{bmatrix} \times dt \tag{3.8}
$$

Isotope effects (α) are input by the user as site-specific fractionation factors, i.e. a separate value for each atomic position. QIRN then creates the rate constants for each isotopologue by multiplying the rate constant of the monoisotopic isotopologue (k_0) by the vector of isotope effects.

$$
\begin{bmatrix} k_0 \\ k_1 \\ \vdots \\ k_{N-1} \end{bmatrix} = \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_{N-1} \end{bmatrix} \times k_0 \tag{3.9}
$$

$$
\alpha_x = \frac{k_x}{k_0} \tag{3.10}
$$

Clumped Isotope Effects

To calculate isotope effects on multiply-substituted isotopologues, QIRN uses the product of the respective site-specific isotope effects, i.e. representing a stochastic (random) effect. Thus by assigning site-specific isotope effects for a molecule, k-values for all of the isotopologues, including those of 'clumped' isotopologues, are automatically defined.

$$
\alpha_{pu} = \alpha_p \times \alpha_u \tag{3.11}
$$

where p and u represent a singly-substituted isotopologues and pu represents their clumped isotopologue. A (non-stochastic) clumped isotope effect represents a deviation from the product rule. Users would implement these effects as γ-values which are multiplied by the appropriate rate constants (defined by Wang et al., 2015). This

feature is not yet available in the GUI but will be in future versions of QIRN.

$$\gamma_{pu} = \frac{\alpha_{pu}}{\alpha_p \times \alpha_u} \tag{3.12}$$

Reversibility and Equilibrium Isotope Effects

QIRN treats reversible reactions as two distinct, unidirectional forward and reverse reactions. QIRN does not include an explicit "reversibility" parameter (eg, Wing and Halevy, 2014) but can achieve the same outcome by varying the ratio of the forward and reverse steady state fluxes. These fluxes are set by the user-defined forward and reverse rate constants for the reaction. For an explanation of defining rate constants see Section 2.7 and the QIRN User Guide.

Equilibrium isotope effects (EIE) are expressed as the ratio between the forward (α_{for}) and reverse (α_{rev}) isotope effects. Thus, if the forward kinetic isotope effect (KIE) is defined, one can define the EIE by assigning the reverse KIE, or vice versa. As described above, EIEs are also assigned separately for each atomic position. If users want reactions to run in the opposite direction of that defined in QIRN's "Reaction Database" (see QIRN User Guide), they can define the reverse flux as being higher than the forward flux through their respective rate constants.

$$\alpha_{eq} = \frac{\alpha_{for}}{\alpha_{rev}} \tag{3.13}$$

Calculating isotopic properties

At the end of the forward model run, QIRN calculates and outputs site-specific isotope compositions (as delta values) for all molecules in the network. It does so by summing the abundances of all isotopologues with a rare isotope at a given atomic site, and dividing this by the of sum of the abundances of all the remaining isotopologues, giving the isotope ratio of the site. This is further converted to a delta value using the reference isotope ratio, and the process is repeated for each site in each molecule. For example to find the carbon isotope composition of the C-1 site in a 3-carbon molecule:

$$^{13}R_{C1} = \frac{I_4 + I_5 + I_6 + I_7}{I_0 + I_1 + I_2 + I_3} = \frac{100 + 101 + 110 + 111}{000 + 001 + 010 + 011} \tag{3.14}$$

$$\delta^{13}C_{C1} = \left(\frac{^{13}R_{C1}}{^{13}R_{ref}} - 1\right) \times 1000 \tag{3.15}$$

To calculate a compound-specific isotope ratio, the isotope ratio of each atomic site can simply be averaged across the number of relevant atoms in the molecule. To arrive at clumped isotope distributions for a molecule or an intramolecular fragment, the sum of isotopologue abundances with the multiple-substitution of interest (i,j) is divided by the sum of the remaining isotopologue abundances. This gives a specific clumped isotope ratio for the molecule or fragment ($^{13}R_{clump, ij}$). A stochastic clumped isotope ratio ($R_{stoch, ij}$) can be calculated from the isotope ratio of the atomic sites involved in the rare isotope clumping, where R_{avg} and z are the average isotope ratio of those sites and the number of sites in that fragment of the molecule, respectively:

$$R_{stoch,ij} = R_{avg,ij}^{z} \qquad (3.16)$$

If the isotope ratios of both sites are known:

$$R_{stoch,ij} = R_i \times R_j \qquad (3.17)$$

$R_{clump, ij}$ and $R_{stoch, ij}$ can be converted to fractional abundance through Equation 4. To calculate the clumped anomaly from the stochastic distribution (Δ):

$$\Delta = (\frac{F_{clump,ij}}{F_{stoch,ij}} - 1) \times 1000 \qquad (3.18)$$

Here we present an example of a doubly-substituted clumped anomaly, but note that the approach is amenable to any number of substitutions.

QIRN Reaction Functions

Reactions involving ≤ 2 reactants and ≤ 2 products have four possible permutations (A to C, A to C+D, A+B to C, and A+B to C+D). QIRN uses different reaction functions for each of these. It also has functions for tracking atomic positions from reactant to product without corrupting their intramolecular isotope distributions. Such complexity must be included since atoms can be rearranged (e.g. C1-C2-C3 molecule reacts to form C2-C1-C3) or exchanged (e.g. C1-C2-C3 + C4 reacts to form C4-C2-C3 + C1) during physical, chemical and biological processes.

The foundation for reliably tracking atoms through reactions is a self-consistent numbering scheme. For applications with carbon isotopes, numbering of each atom in QIRN is based mostly on IUPAC naming (or is otherwise explicitly specified) to maintain self-consistency throughout the reaction network. For each of the other elements (N, S, H, etc.), a database should be created to record the assigned
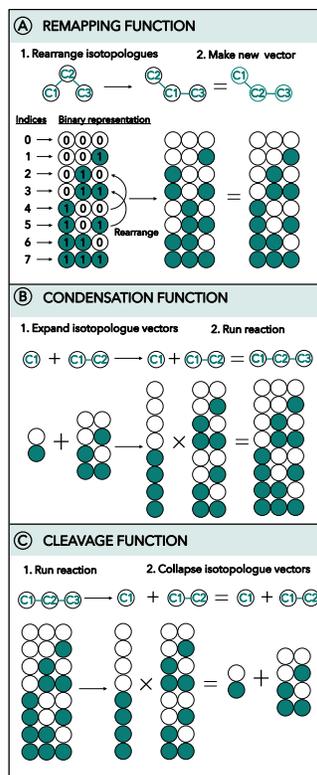
Figure 3.1: Reaction functions built into QIRN to track the isotopic properties of molecules through inter- or intramolecular atomic rearrangement.

atomic ordering of compounds as they are needed. These databases can then be referenced by future models for self-consistency. Nevertheless, even with carbon isotopes, some numbering problems still arise, mostly when functional groups cause changes in the direction of IUPAC numbering. For example, the reaction of dihydroxyacetone (DHAP) to glyceraldehyde-3-phosphate (G3P) involves no carbon-carbon bond cleavages or formations, yet the C-1 site of DHAP becomes the C-3 site of G3P based on IUPAC naming. Without rearrangement of the isotopologue array, QIRN would confuse the C-1 carbon atoms of DHAP and G3P. QIRN uses different combinations of the following four functions to maintain consistent atomic ordering and track intramolecular rearrangement or intermolecular exchange:

Transformation Function

In the simplest example, a molecule undergoes a reaction in which no atoms are rearranged or exchanged and the IUPAC numbering is stable, e.g. the reaction of pyruvate to alanine while tracking C isotopes. In this case, the isotopologue distribution follows Equation 8 without further manipulations.

Remapping Function

If carbon atoms are rearranged over the course of a reaction, or the IUPAC numbering of reactant and product do not match, then the isotopologue distribution must be artificially manipulated to ensure the accurate transfer of isotope composition of each atomic site from reactant to product. After the reaction step (Equation 8) is calculated, QIRN remaps the array indices of the product isotopologues to match the product's intended carbon ordering. As a concrete example (Figure 3.1A), if the C-1 and C-2 sites of a C_3 reactant were switched during a reaction, QIRN would need to swap indices 010 with 100 and 011 with 101 to maintain order. Each reaction that requires atomic rearrangements is given a barcode (represented by a string of integers) that QIRN references when the reaction is run. Rearrangement barcodes are user assigned within the Reaction Database. Each reaction's barcode is assigned once and can thereafter be referenced in new networks. This reaction barcode can also be used to treat molecules that are symmetrical. When a reactant is symmetrical, QIRN rearranges the isotopologue array to average the abundances of redundant isotopologues, i.e. those with isotopic substitutions in symmetrically equivalent positions. This occurs during the reaction itself rather than on the total pool of reactant, since the same substrate may not be symmetrical to every reaction, as is the case for prochiral enzymes (Ochs and Talele, 2020).

Condensation Function

When two substrates combine to form a larger molecule, the condensation function is used. QIRN first expands the reactant isotopologue vectors (abundances and rate constants) to the length 2^M where M is the total number of atoms of interest in the product. If two products are formed, M is the sum of the atoms in those products. Repeating units of the reactants' isotopologue arrays populate these expanded vectors and represent the full range of product isotopologues. The reactant isotopologue abundances are then divided by the number of repeating units necessary to populate the expanded (2^M) arrays. For example in the case of a reaction between a C_2 reactant (A) and C_1 reactant (B), the reactant arrays expand to form the C_3 product (D):

$$
\begin{bmatrix} I_{A0} \\ I_{A1} \\ I_{A2} \\ I_{A3} \end{bmatrix} + \begin{bmatrix} I_{B0} \\ I_{B1} \end{bmatrix} \implies \begin{bmatrix} I_{A0} \times 0.5 \\ I_{A1} \times 0.5 \\ I_{A2} \times 0.5 \\ I_{A3} \times 0.5 \\ I_{A0} \times 0.5 \\ I_{A1} \times 0.5 \\ I_{A2} \times 0.5 \\ I_{A3} \times 0.5 \end{bmatrix} \times \begin{bmatrix} I_{B0} \times 0.25 \\ I_{B0} \times 0.25 \\ I_{B0} \times 0.25 \\ I_{B0} \times 0.25 \\ I_{B1} \times 0.25 \\ I_{B1} \times 0.25 \\ I_{B1} \times 0.25 \\ I_{B1} \times 0.25 \end{bmatrix} \times \begin{bmatrix} k_0 \\ k_1 \\ k_2 \\ k_3 \\ k_4 \\ k_5 \\ k_6 \\ k_7 \end{bmatrix} \times dt = \begin{bmatrix} I_{D0} \\ I_{D1} \\ I_{D2} \\ I_{D3} \\ I_{D4} \\ I_{D5} \\ I_{D6} \\ I_{D7} \end{bmatrix} \qquad (3.19)
$$

The two expanded reactant arrays are multiplied together and by the second-order reaction rate constants to calculate the flux of the condensation reaction (Figure 3.1B). This procedure assumes the reactant and product condense at the site between the last atom of Reactant A and the first atom of Reactant B. If this is not the case (i.e. acetyl-CoA adding to the middle of oxaloacetate's structure to synthesize citrate), a remapping step takes place on the product array.

Cleavage Function

The reverse of the condensation function involves the cleavage of one molecule into two. Once the reaction has run, QIRN simplifies the products' isotopologue abundance vectors into two shorter vectors by summing degenerate isotopologues (Figure 3.1C). If two substrates react to form two products of differing lengths, the condensation and cleavage functions occur consecutively with a remapping function in between.

Flux Balancing to Steady State

All QIRN models reported here were run out to isotopic steady state. In order to achieve steady state, the source and sink fluxes for each compound in the network must be balanced. But not all combinations of rate constants will necessarily reach steady state. They must be balanced by the user to achieve that goal. In some cases, all substrates in the reaction network were initially set to a concentration of one. Fluxes are equivalent to user-defined rate constants (Equation 5). In these scenarios, reaction rate constants were fully constrained to a set that represented the branching points of the network while making the input and output of each substrate equivalent. These fluxes were, in this study, taken or deduced (as described in the Section 2.8) from previous work and then directly input into QIRN as rate constants.

Since source and sink reaction rates at each substrate node were equivalent, the concentration of substrates did not change over time. However, the initial isotopic composition of these substrates were not equal to their final values. They changed with time until they reached isotopic steady state. The QIRN GUI allows the user to visualize the compound-specific isotopic composition of every substrate in the network over time to determine whether the model has reached isotopic steady-state.

In other cases, only a subset of the substrates were given initial concentrations and the downstream products were filled as the reaction network progressed through time. In these simulations, substrate concentrations changed with time and reactions' rate constants did not have a simple relationship with their steady state absolute fluxes. To find the appropriate reaction rate constants that matched the desired fluxes, an optimization algorithm was written into QIRN that inverts rate constants from user-assigned reaction fluxes. Here, the user chooses steady state fluxes and QIRN uses an ordinary-differential-equation (ode) integrator and optimization Python package (scipy) to find reaction rates that best fits the fluxes. We found this approach most useful when networks involved branching points with three outputs with equivalent fluxes (e.g. Calvin Cycle). In these scenarios, 33.3% of the input flux goes to each output reaction so that the overall output is only 99.9%. Even when outputs were assigned with more significant figures, the deviation between input and output fluxes at these nodes eventually caused network instabilities that drove the networks out of steady state. Using this inversion algorithm, QIRN found reaction rate sets that satisfied the users desired fluxes without creating network instabilities. The deviations between the user-defined and optimized flux sets are reported. The 'Reaction Rate Conversion' module is accessible in the GUI. For more information, see the Supplementary Materials.

Model Parameters

***Zymomonas mobilis* metabolism**

An absolute flux map of *Zymomonas mobilis*' central metabolism was obtained from the work of Jacobson et al., 2019. These fluxes were determined by growing different cultures of *Z. mobilis* on multiple forms of ^{13}C-labelled glucose including those labelled at the C-3 and C-6 sites. The relative distribution of monoisotopic (M), singly substituted (M+1) and doubly substituted (M+2) forms of central metabolites constrained the central metabolic fluxes of *Z. mobilis*'s enzymes (Jacobson et al., 2019). Rate constants were defined in QIRN to yield those same metabolic fluxes.

A boundary condition of constant glucose concentration was specified. In two different simulations, glucose was separately labelled (99%) at the C-3 and C-6 sites (i.e. only one ^{13}C per molecule). For the reaction between phosphoenolpyruvate and bicarbonate to form oxaloacetate, bicarbonate was set as a constant concentration (i.e., infinite reservoir) with a δ ^{13}C value of 0‰. All other metabolites were set to an initial concentration of 1 and a δ^{13}C of 0‰ for all carbon sites. Both simulations were run to isotopic steady state (500,000 timesteps). Such long simulation times were required because the offset in isotopic compositions of glucose and that of the downstream products was initially very large. The ^{13}C labeling patterns of thirteen metabolites were compared between *in vivo* experiments by Jacobson et al. and QIRN simulations (this study): Glucose-6-phosphate, 2-keto-3-deoxy-6-phosphogluconate, glyceraldehyde-3-phosphate, pyruvate, alphaketoglutarate, succinate, malate, citrate, fumarate, aspartic acid (proxy for oxaloacetate), ribulose-5-phosphate, dihydroxyacetone, and phosphoenolpyruvate. Full descriptions of the parameters used for the *Z. mobilis* model can be found in the QIRN "Network Builder" input files supplied in the Supplementary Materials.

Strecker Synthesis

The site-specific carbon isotope composition of alanine in the meteorite Murchison was previously measured and interpreted using a steady-state model of Strecker synthesis that is presumed to occur on meteorite surfaces (Chimiak et al., 2021). Here we recreated this model in QIRN as a test of site-specific predictions. Two branching points exist in the network (Figure 3.3), both with 50% branching ratios. The reversibility of Reactions 3 and 5 were assumed to be 50% and 80%, respectively. The starting isotopic compositions of carbon monoxide (CO), methyl moieties (CH$_3$), and cyanide were 166‰, -36‰ and -7‰, respectively, as described by Chimiak et al., 2021. An extra reaction was added to the network to remove alanine from the system. This was necessary to report the isotope ratio of alanine being produced at steady-state, rather than its integrated composition over the entire forward model. For further discussion see Section 4.4.3. Full descriptions of the parameters used for the Strecker synthesis model can be found in the QIRN "Network Builder" input files supplied in the Supplementary Materials.

Microbial Sulfate Reduction

Wing and Halevy Wing and Halevy, 2014 built a steady-state metabolic model of the dissimilatory sulfate reduction pathway involving one step of physical diffusion and three enzymatic reactions. In this model, the reversibility of each step determined the net ^{34}S fractionation between sulfate and sulfide. These were calculated from exogenous environmental parameters which act as inputs into the model. Here, QIRN was compared to the Wing & Halevy model using a Matlab implementation of the latter built by Sim et al., 2019 - used in this study with permission of the authors.

Extracellular sulfate concentration and cell-specific sulfate reduction rate (cSRRr) were changed independently within the Wing & Halevy model. Sulfate concentration was either 2 or 28 mM. cSRRr was varied between 0 and 175 fmol/cell/day. Temperature and extracellular sulfide were held constant at 25°C and 1 mM, respectively. Equilibrium and kinetic isotope effects on each reaction can be found in the Table S1. Each combination of sulfate concentration and cSRRr generated reversibility terms for each reaction and net ^{34}S-fractionation between sulfate and sulfide. QIRN sampled this parameter space ten times in each regime of sulfate concentration.

In QIRN, the net flux and concentration of extracellular sulfate were 0.005 and 1 (unitless), respectively. Sulfate concentration was set as a constant boundary condition with a $\delta\,^{34}$S value of 0‰. QIRN's net flux and sulfate concentration were arbitrary and did not impact the overall fractionation at steady state, because the reversibility terms were the only control over the net fractionation (see Wing and Halevy, 2014 Equation 1). The forward and reverse fluxes for each reaction in QIRN were calculated from QIRN's net flux and reversibility terms calculated by the Wing & Halevy model (Equations 19 and 20).

$$J_{rev} = \frac{J_{net}}{1/f - 1} \tag{3.20}$$

$$J_{for} = J_{net} - J_{rev} \tag{3.21}$$

where J_{net}, J_{rev}, and J_{for} are the net, reverse, and forward fluxes, and f is the reversibility (f = J_{rev}/J_{for}). J_{net} was set by the user in QIRN and f was calculated by the Wing & Halevy model with the MATLAB script provided by Sim et al.,

2019. The overall fractionation was calculated by comparing steady state $\delta^{34}S$ of intracellular sulfide and extracellular sulfate:

$$^{34}\alpha = \frac{\delta^{34}S_{sulfate} + 1}{\delta^{34}S_{sulfide} + 1} \tag{3.22}$$

$$^{34}\epsilon = {}^{34}\alpha - 1 \tag{3.23}$$

All simulations were run for 30,000 timesteps to achieve isotopic steady state. Full descriptions of the parameters used for the sulfate reduction model can be found in the QIRN "Network Builder" input files supplied in the Supplementary Materials.

Calvin-Benson-Bassham Cycle

The Calvin-Benson-Bassham (CBB) cycle in plants involves over 20 reactions and 3 physical transport processes, and serves as a good test for QIRN's ability to handle more complicated networks. The system was previously modeled at steady state by Tcherkez et al., 2004 based on the site-specific carbon isotope composition of plant glucose (Rossmann, Butzenlechner, and H.-L. Schmidt, 1991). In the Tcherkez et al. model, fluxes were calculated through a system of equations with two input parameters: 1) export rate of fructose-6-phosphate from the chloroplast, and 2) the ratio of carboxylation to photorespiration in the enzyme ribulose-5-bisphosphate carboxylase/oxygenase (RuBisCo). These fluxes were normalized to unity with the rate of carboxylation being one. In this previous study, Tcherkez et al. reported multiple regimes of photorespiration, changing the isotope effects of enzymes within the CBB cycle based on the photorespiration rate to match measured glucose intramolecular $\delta^{13}C$ values from Rossmann, Butzenlechner, and H.-L. Schmidt, 1991. QIRN reconstructed this model with increasing photorespiration between 9% and 47% total activity of RuBisCo while keeping isotope effects constant to assess the affect of photorespiration on the carbon isotope structure of glucose. Fructose-6-phosphate export was held at 0.056 (unitless). This export was held constant to reproduce the Tcherkez model. All other fluxes changed in response to maintain steady state. Other than the isotope effects (which are held constant here), every parameter in the

Tcherkez model was matched here in QIRN. The initial concentrations of ribulose-5-phosphate and carbon dioxide were 10 and 1 with δ ^{13}C values of -20‰ and -8‰, respectively. CO_2 was assigned a constant concentration. All other metabolites began with a concentration of zero. Simulations were run for 20,000 time steps. Isotope effects were applied on both transketolase enzymes, fructose-6-phosphate aldolase, RuBisCo, and glycine decarboxylase (see Table S2). These effects were taken from the 23% photorespiration regime as deduced by Tcherkez et al.. All reactions were assumed to be irreversible. The same substrate in multiple compartments are treated as separate species (i.e. "G3P" and "G3P-cytosol"). Physical transport processes between compartments are treated as reactions that consume a molecule in the source compartment and create the same molecule in the product compartment. For further details on fluxes and fractionation factors, see the QIRN input "Network Builder" files for this case study in the Supplementary Materials.

Fatty Acid Biosynthesis

Monson and Hayes (1980) measured and interpreted the intramolecular carbon isotopic structure of fatty acids synthesized by *Escherichia coli* when it was grown on glucose (Monson and J. M. Hayes, 1980). To update their model of *E. coli* metabolism to the contemporary understanding, a metabolic flux model of *E. coli* grown on glucose under similar conditions by Wijker et al. (Wijker et al., 2019) was coupled to the fatty acid biosynthesis model explored by Monson & Hayes. Glucose was given a -10‰ compound-specific $\delta^{13}C$ value with no intramolecular heterogeneity as is assumed in the original model. Both bicarbonate (0‰) and glucose were set as constant-concentration boundary conditions. All of the fluxes estimated by Wijker et al., 2019 were implemented into QIRN as well as several reactions not explicitly stated in their model but required to complete the metabolism of *E. coli*, including export of the amino acids glutamate, aspartate, alanine/valine, serine, and aromatic amino acids from alpha-ketoglutarate, oxaloacetate, pyruvate, 3-phosphoglycerate, and ribose-5-phosphate, respectively. The fluxes of these reactions were deduced from the Wijker et al. model by balancing the source and sink reactions of metabolites to maintain steady state. Fractionation factors were only applied to carboxylating or decarboxylating reactions for simplicity (Table S3). Notably, pyruvate dehydrogenase was assigned 20‰, 23.2‰, and 3‰ ^{13}C-fractionations on the C-1, C-2, and C-3 positions of pyruvate, respectively, as was measured by DeNiro and Epstein, 1977. For further details on fluxes and fractionation factors, see the QIRN input "Network Builder" files for this case study in the

Supplementary Materials.

Acetyl-CoA is polymerized into fatty acid molecules through the malonyl-CoA elongation pathway. However, rather than building C_{14}, C_{16} and C_{18} fatty acids (>300,000 total isotopologues), we approximated these molecules as as C_4, C_6 and C_8 fatty acids. When compound-specific $\delta^{13}C$ values were calculated, the first ten carbon atoms on each acid were assumed to have the same intramolecular $\delta^{13}C$ as five repeating units of acetyl-CoA. The first ten atoms of a fatty acid are created by unidirectional polymerization of acetyl-CoA. Thus, no isotope effect could be expressed on these first ten carbon atoms and they would inherit their isotopic composition from acetyl-CoA. From the C_{14} to C_{18} fatty acids, branching points occur in the elongation pathway. Fatty acids either continue to elongate or are hydrolyzed out of the biosynthetic pathway to be used in the cell membrane. Branches in the elongation pathway were determined by the relative abundance of fatty acids measured in *E. coli* cells. A C-1 site-specific isotope effect (14‰, adopted from Monson and J. M. Hayes, 1980) was applied to all hydrolysis steps that removed fatty acids from the elongation pathway. Forward models were run for 30000 timesteps. All reactions were assumed to be irreversible.

QIRN GUI: Inputs and data visualization

To run a simulation in QIRN, the user must know (or assume) the forward and reverse rate constant of every reaction in the network, as well as any position-specific isotope effects on those reactions. These can be input into a parameter file (csv format) known as the "Network Builder" file. The user must also choose the starting conditions (concentrations and isotopic compositions) of the substrates and whether these will act as free substrates with changing compositions or as fixed boundary conditions. This information is input into the second parameter file known as the "Intermediates Database". These two parameter files are then passed to QIRN through the GUI. The GUI asks the user how long (in seconds, with a 0.1 second timestep) to run the forward model. After the model has run, the user can work through the GUI to explore plots of reaction fluxes, substrate concentrations, and isotope compositions with time. Using these plots, the user can determine whether the network has reached steady state, at which point they can evaluate final isotopic compositions. In the GUI module, the site-specific isotopic compositions can be reported for any substrate in the network. The user can also export the raw QIRN data for visualizations and other functions not available in the GUI. More detailed instructions on how to use the input parameter files and the GUI are in a full QIRN

User Guide in the Supplementary Materials.

3.4 Results

Zymomonas mobilis metabolism

The distribution of ^{13}C label in thirteen intracellular metabolites measured from *in vivo* cultures of *Z. mobilis* are compared against those produced by simulations in QIRN (Figure 3.2). *Z. mobilis* was grown under two conditions (C-3 ^{13}C and C-6 ^{13}C labelling). Under both conditions, every metabolite measurement is represented by three data points (M, M+1, and M+2), resulting in 78 total comparisons. The root-mean square deviation (RMSD) between model and measurements is 4.9%. (Note: percent not permille, because of the high level of ^{13}C labeling.) Measurement uncertainty on *in vivo* data is typically between 1-2% (1 SD) (Jacobson et al., 2019). There are four data points from the C6-labelling simulations that are offset from the *in vivo* values by >14%, heavily skewing the overall RMSD (open circles, Figure 3.2). These points represent the M and M+1 values of fumarate and malate. Malate also has the highest offset in the C3-labelling simulation with a residual of 6% for both M and M+1 relative abundances. In the Jacobson et al. model, fumarate and malate are synthesized through unidentified biosynthetic pathways, which is problematic for our model. Here, we generated them directly from oxaloacetate. If the six fumarate and malate data points are excluded from the comparison, the RMSD is reduced to 1.3%, within error of the uncertainty on *in vivo* measurements.

Strecker Synthesis

The Chimiak et al (2020) model of amino acid Strecker synthesis from CO, CH$_3$ and cyanide is shown in Figure 3.3A. Applying fractionations to Reactions 3 and 5, QIRN accurately reproduced the intramolecular δ^{13}C values of alanine measured from the meteorite. In this network, CO became the C-2 position of alanine. Indeed, the extreme ^{13}C enrichment from CO did end up at the C-2 position of alanine (Figure 3.3B). Comparing the ^{13}C/^{12}C ratio at all three sites between QIRN and the meteoritic values, there was an RMSD of 4.7‰, well within error of experimental uncertainty (Chimiak et al. 2020). We note that the previous model used the measured values to invert the network's isotope effects, thus reproducing meteoritic compositions is equivalent to showing that QIRN accurately reproduced the previous model. To further demonstrate the suitability of QIRN isotopic calculations, the C-1 site-specific fractionation on the hydrolysis of alpha-aminonitrile to alanine was allowed to vary between 5 and 30‰ (Figure 3.3C). As expected, the differ-
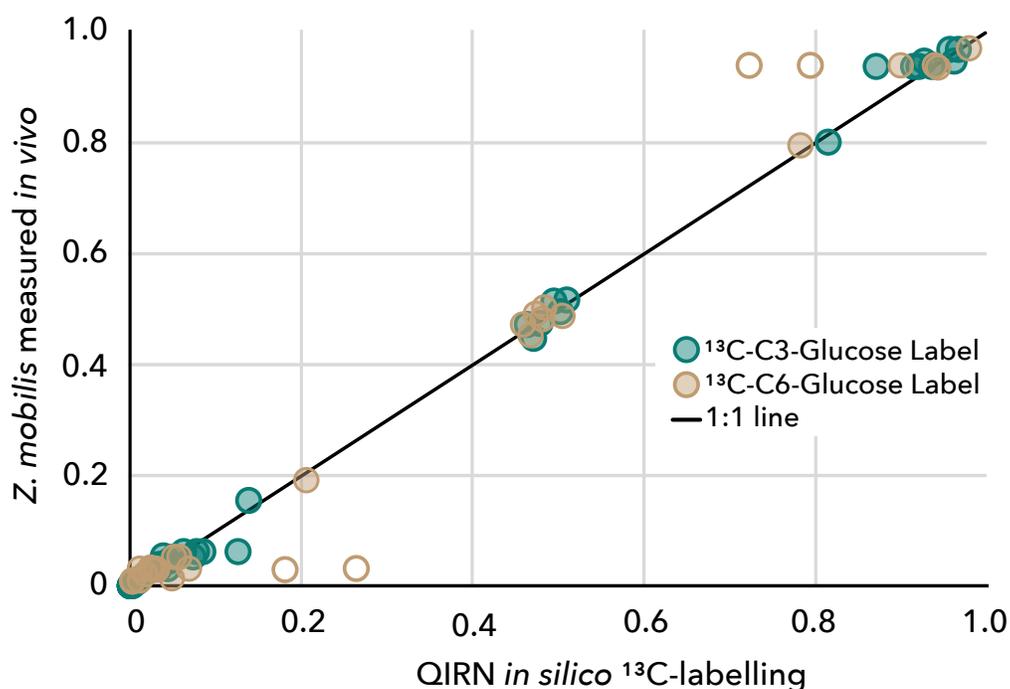
Figure 3.2: The bacterium *Zymomonas mobilis* was grown with glucose ^{13}C-labelled at either the C3 or C6 site. From the distribution of M, M+1, and M+2 labelling on 13 central metabolites an absolute flux map of *Z. mobilis*'s metabolism was constructed by Jacobson et al., 2019. This network was reconstructed in QIRN, with fluxes matching those of Jacobson et al. QIRN simulations in which glucose was *in silico* labelled at its C-3 and C-6 sites produced metabolites with ^{13}C-labelling that matches those observed *in vivo* with an RMSD of 4.9%. Each metabolite is represented by three data points for M, M+1 and M+2 labelling. Open circles represent anomalous data points from malate and fumarate. Abbreviations: M = monoisotopic, M+1 = single ^{13}C substitution, M+2 = double ^{13}C substitution.

ence between the δ^{13}C values of C-1 in alanine and alpha-aminonitrile changed commensurately with the applied fractionation.

QIRN can also calculate isotope ratios of clumped species. Although no clumped isotope fractionations were specified for this network, ^{13}C-^{13}C anomalies from the stochastic isotopologue distribution can still occur due to combinatorial effects. When the abundance of ^{13}C-^{13}C clumping between C1-C2 and C2-C3 of alanine was calculated, both were significantly less than the stochastic distribution. This anomaly became more depleted by 7‰ as the C-2 site of alanine was artificially increased from 100 to 200‰ (Figure 3.3D). The clumped anamoly on the C2-C3 sites was less severe because the difference in the δ^{13}C values between those sites

was smaller.
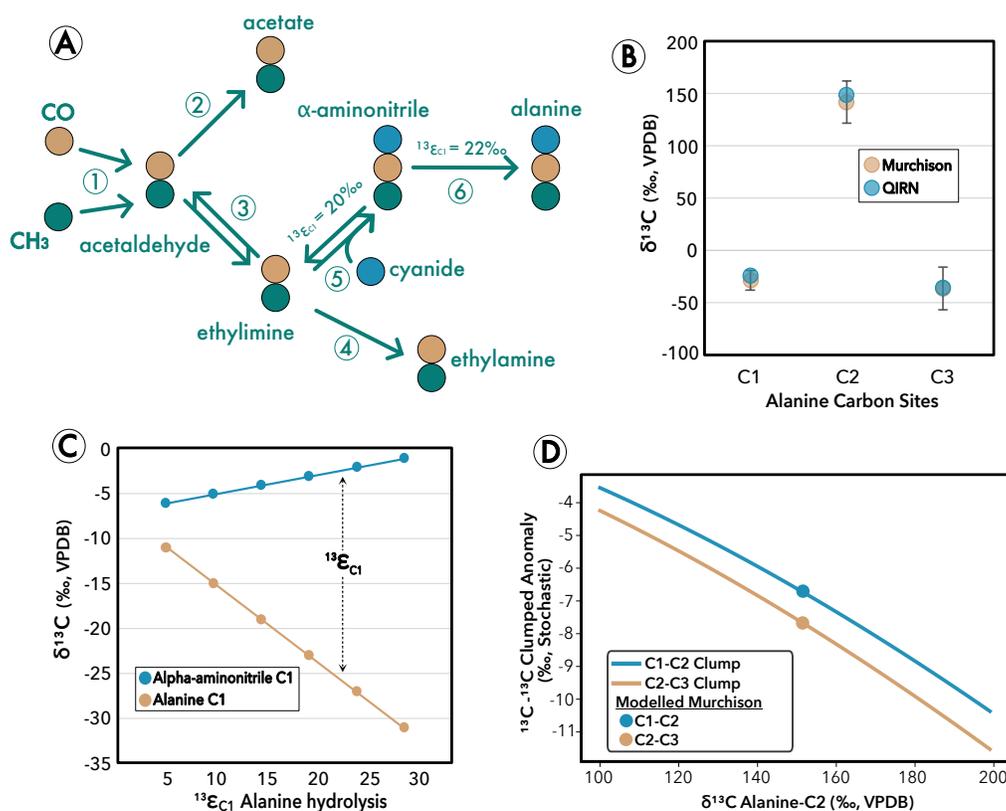
Figure 3.3: Meteoritic alanine intramolecular $\delta^{13}C$ distribution is predicted to arise from Strecker synthesis. QIRN reproduces this model of alanine synthesis within error of measurements made on alanine from the Murchison meteorite. (A) was adapted from Chimiak et al., 2021.

Sulfate Reduction

Established models of ^{34}S fractionation ($^{34}\epsilon$) between sulfate and bisulfide during dissimilatory microbial sulfate reduction (MSR) demonstrate a nonlinear, monotonic decrease in $^{34}\epsilon$ with increasing cSSRr (dotted lines, Figure 3.4). This trend is reproduced by QIRN across ten values of cSSRr between 1 and 175 fmol/cell/day. The relationship between cSSRr and $^{34}\epsilon$ can be offset by lowering the extracellular sulfate concentration from 28 mM to 2 mM within the Wing & Halevy model, causing the overall fractionation to decrease. QIRN matches the model outputs in both sulfate concentration regimes, accurately reproducing the nonlinear trends across a two-dimensional parameter space. Root mean square deviation (RMSD) between $^{34}\epsilon$ values from the Wing & Halevy model and QIRN's outputs were 0.09‰ and 0.05‰ for 28 mM and 2mM sulfate concentration regimes, respectively.
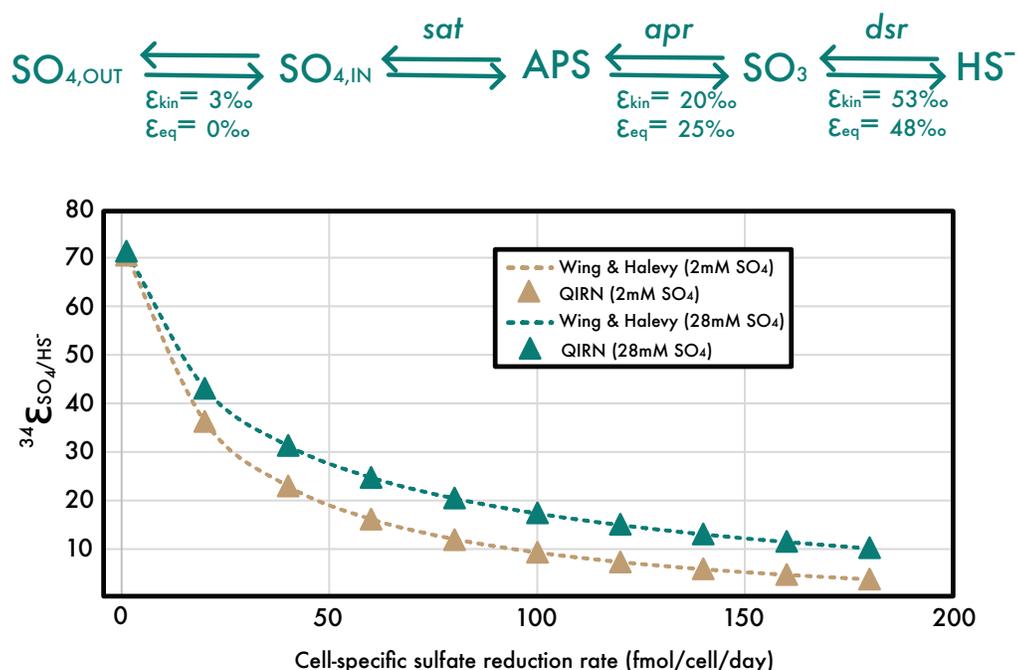
Figure 3.4: QIRN coupled to a metabolic model of sulfate reduction reproduces net sulfur isotope fractionations across changing cell-specific sulfate reduction rate (cSSRr) and sulfate concentration.

Calvin-Benson-Bassham (CBB) Cycle

Steady-state models of isotope fractionation in the CBB cycle of plants capture much of the variance in the intramolecular $\delta^{13}C$ distribution of the six-carbon sugar glucose as described in Tcherkez et al. (2004). QIRN accurately reproduces these results. At a photorespiration rate (PR) of 23% the offset between the six glucose carbon sites' carbon isotope compositions predicted by Tcherkez et al. model outputs and QIRN was 0.66‰ (RMSD). Except for the C-6 site, these models accurately predict the site-specific anomalies in glucose from sugar beet syrup and maize corn flour (data from (Rossmann, Butzenlechner, and H.-L. Schmidt, 1991)). Using QIRN, PR was varied between 9% and 47% (of the total CO_2 input flux; Figure 3.5) and resulting changes in glucose intramolecular isotope composition were recorded. QIRN predicts a significant exaggeration of glucose's intramolecular isotopic variability as photorespiration increases, particularly at C-4. The C-4 site reached an enrichment of almost 9‰ relative to the molecular average carbon isotope composition when PR rate was 47%.
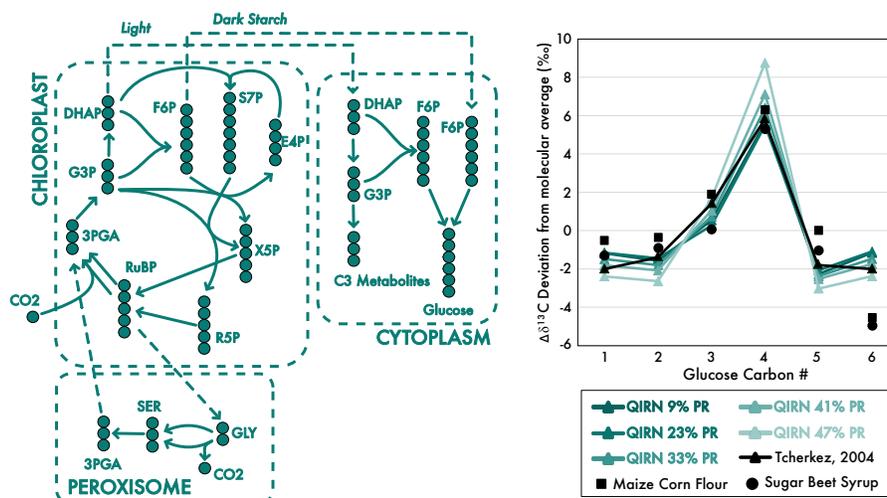
Figure 3.5: Glucose intramolecular $\delta^{13}C$ structure measured in sugar beet syrup and maize flour are reproduced by coupling a isotopic model of photosynthesis in plants by Tcherkez et al., 2004 QIRN can then explore multiple regimes of photorespiration rate to explore the effect on glucose position-specific carbon isotope composition.

Fatty Acid Biosynthesis

A model of *E. coli*'s central metabolism and another of fatty acid biosynthesis (Figure 3.6A) were coupled using QIRN to predict the fatty acid intramolecular $\delta^{13}C$ distributions. Acetyl-CoA created by the model of central metabolism had a bulk δ ^{13}C value of -12.25‰ (Figure 3.6B). The difference between the C-1 and C-2 sites of this molecule was 3.6‰. QIRN output matched the relative abundance of fatty acids from *E.coli* cells (Figure 3.6A). These were defined in QIRN with the fluxes through the enzyme-catalyzed hydrolysis reactions producing lipids (Figure 3.6). The C-1 position of the fatty acids in QIRN simulations were all ^{13}C-depleted compared to their bulk values, but by different amounts. QIRN reproduced the measured C-1 $\delta^{13}C$ values from *E. coli* cells with an RMSD of 0.9‰ (Figure 3.6C). The pattern of compound-specific $\delta^{13}C$ values of the lipids was also reproduced by QIRN with an RMSD of 0.5‰. (Figure 3.6D). This was achieved without any tuning of the site-specific isotope effects in *E. coli*'s central metabolism or lipid biosynthetic pathway. Only published, measured or deduced literature values of isotope effects and metabolic fluxes were implemented.
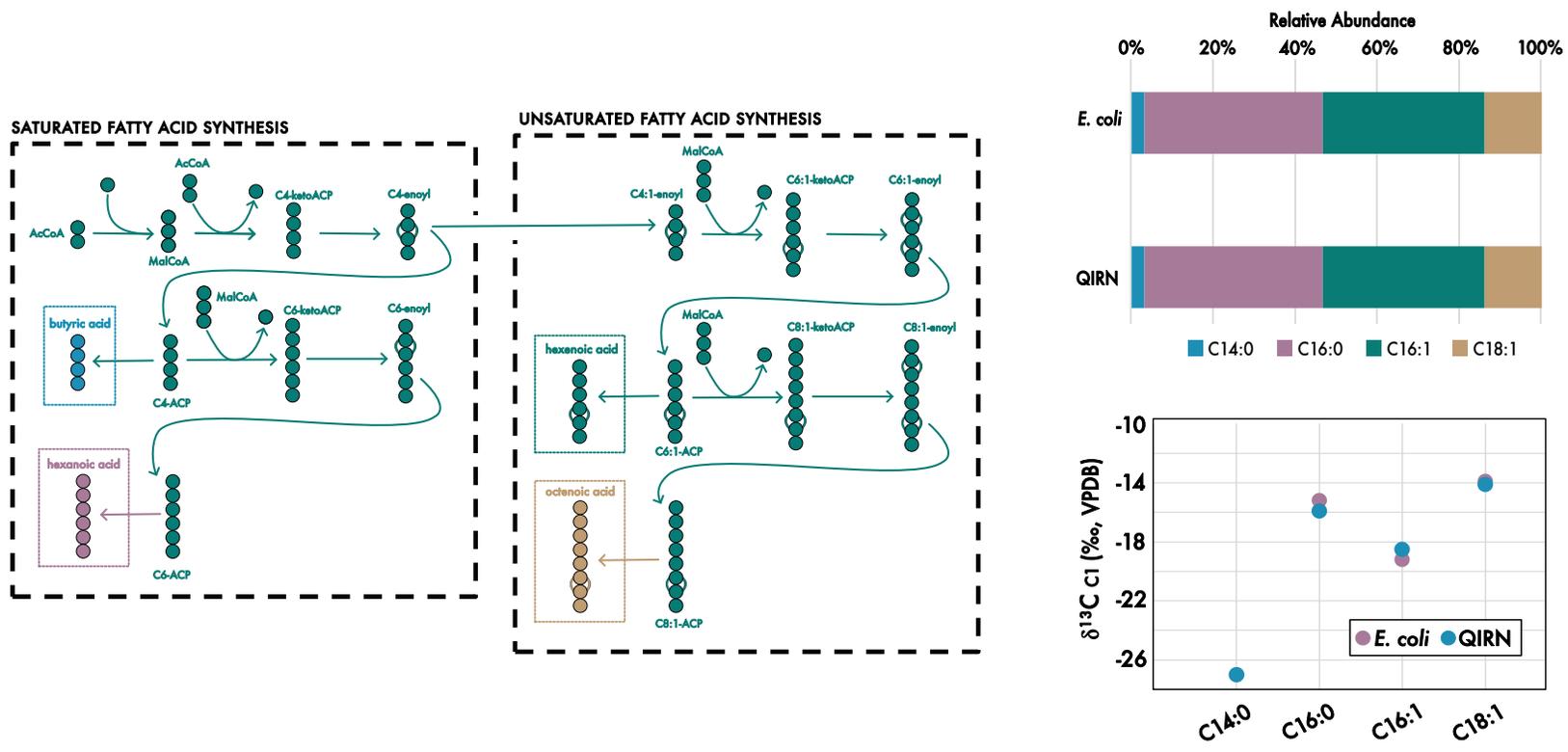
Figure 3.6: Site-specific $\delta^{13}C$ at the C1 position of fatty acids produced by *E. coli* can be explained by applying a fractionation factor on the enzymatic hydrolysis of acyl-ACP intermediates as demonstrated by Monson and J. M. Hayes, 1980. QIRN couples this model to a larger metabolic model of *E. coli* grown on glucose and matches (A) the relative abundance, (B) C1 site-specific $\delta^{13}C$, and (C) molecular-average $\delta^{13}C$ of fatty acids found in *E. coli*.

Computational Expense

To explore the relationship between reaction network size and computational expense, the four reaction networks described above were run on a laptop computer for 10,000 steps and timed, each in triplicate. Standard deviations of run time were all <1 sec. All simulations were performed in the same session without making any changes to the program. Despite the high complexity of the *E. coli* central metabolism model, 10^4 timesteps ran in <15 sec. As reactions became more simplistic (fewer reactions and molecular remapping steps), computational expense decreased. The other metabolic pathways representing the CBB cycle and *Z. mobilis* fermentation were completed in less than 10 seconds. Correlation between computational time and number of reactions was monotonic and nominally linear except for the smallest network. MSR took less than half-a-second, much faster than the similarly sized Strecker network. Precise reasons for this difference are not currently known (see Discussion).

3.5 Discussion

QIRN represents an unconventional approach to modelling isotopes at natural abundance in reaction networks via the explicit tracking of isotopologues. This is a closer approximation of how isotope effects are manifest at the molecular level than are "fractionation factors", so the approach is theoretically sound. However, there remain numerous questions about the accuracy, generalizability, flexibility, and computational expense of our particular implementation. We are unaware of any datasets with sufficient natural abundance, position-specific isotope measurements to provide a robust validation of QIRN, so instead we turn to replicating the results of other existing models that use the more conventional, intrinsic-property approach. The following sections describe the results of those comparisons. In replicating these previous models with QIRN, we are also able to make several new observations and predictions about the studied systems.

Proof of concept

Atomic tracking in molecular rearrangements

A first requirement of the model is that it accurately tracks atoms through reactions, rearrangements, and numbering changes. To demonstrate this, our first QIRN model targeted the metabolism of *Zymomonas mobilis*, an obligately fermenting bacterium that uses the Entner-Doudoroff pathway to quantitatively convert glucose into ethanol. This pathway involves complex intramolecular rearrangements of

central metabolites. We reconstructed *Z. mobilis*'s central metabolism in QIRN without specifying any isotopic fractionations to test whether its reaction functions accurately tracked isotopic labels. Indeed, QIRN reproduced the *in vivo* ^{13}C label measurements of thirteen metabolites within analytical uncertainty (Figure 3.2). Modelled labelling of fumarate and malate were outside of error; however, this can be explained by *Z. mobilis*'s truncated TCA cycle, which requires an enigmatic anaplerosis of fumarate and malate through undetermined biosynthetic pathways (Jacobson et al., 2019). In QIRN and the Jacobson et al. model, fumarate and malate are supplied directly from oxaloacetate instead. Since this pathway is not representative of the actual biosynthesis, it is not surprising that the labelling pattern produced by QIRN would differ from the measured one. Thus, despite many molecular rearrangements, QIRN accurately tracked ^{13}C-labelling in *Z. mobilis* metabolism, reproducing *in vivo* isotopic signals of central metabolites from a model bacterium. While this network does not represent the entirety of *Z. mobilis*'s metabolism, it does reflect the contemporary standard for metabolic modeling. Thus, QIRN's reaction functions are reliable and have enough computationally efficiency to simulate a high level of network complexity.

Isotopic fractionations

A second requirement is that QIRN must also impose and accurately calculate isotope effects (and the resulting fractionations) within networks. To test this, isotope effects were implemented into a simple model of abiotic alanine synthesis (Figure 3.3). This Strecker synthesis network is thought to occur on meteorites and lead to the abiotic synthesis of α-amino acids. The intrinsic-property model of ^{13}C-fractionations during Strecker synthesis was originally formulated by Chimiak et al. based on measurements of alanine from the Murchison carbonaceous chondrite. The C-2 position of meteoritic alanine had an extreme ^{13}C-enrichment that was thought to be inherited from a highly ^{13}C-enriched carbon monoxide moiety. That CO moiety in turn begins as the C-1 position of acetaldehyde but is ultimately rearranged to become the C-2 position of alpha-aminonitrile and alanine. Using the molecular remapping function in QIRN, this ^{13}C-enriched carbon site was accurately modelled through the Strecker synthesis reaction network (Figure 3.3A and 3.3B). The validity of QIRN's isotopologue tracking approach was further demonstrated on the C-1 site, whose $\delta\ ^{13}$C value had a predictable relationship with the isotope effect on Reaction 6 (Figure 3.3C). These results confirm that assigning variable rate constants on individual isotopologues is an effective method for implementing

fractionation factors within reaction networks and– when coupled to QIRN's reaction functions – can accurately represent position-specific isotope fractionations in steady state networks.

Generalizability

A major advantage of QIRN's approach to modelling stable isotope compositions is its multifaceted generalizability. By treating isotopologues as individual species in a large finite-difference model, QIRN can autonomously build reaction networks and come to steady state numerical solutions. Changes in topology are simple to implement and can expand into complex networks that, to date, have exceeded 80 reactions and 11,000 isotopologues. For example, if the user wanted to determine the importance of the branching point between Reactions 2 and 3 in Figure 3.3A, we could eliminate Reaction 2 by deleting its row in the Network Builder file and rerunning the simulation. If the user wanted to consider the impact of Strecker synthesis on alanine's isotope composition in a biological system, they could simply copy the Network Builder rows from the Strecker synthesis pathway into the *E. coli* metabolic Network Builder file. Moreover, every atomic position and molecular fragment within a given binary isotope system is represented in the same simulation. Thus, any changes to network topology are imposed on every isotopic property at once. In the following sections, we demonstrate QIRN's generalizability and flexibility using Strecker synthesis networks and sulfur isotope fractionation during MSR as examples.

Reversibility

Reaction reversibility is an important concept that often controls the stable isotope compositions of molecules within a reaction network, as it determines the predominance of kinetic versus equilibrium isotope effects. In QIRN, any reaction can be reversible or irreversible. The extent of each reaction's reversibility is user-controlled through the ratio of its forward and reverse fluxes. This can be adjusted without affecting the net flux of that reaction (Equation 20). To demonstrate the utility and flexibility of this approach, a model of ^{34}S-fractionation ($^{34}\epsilon_{SO4/HS^-}$) during sulfate reduction to sulfide was reproduced using QIRN.

Here, QIRN reconstructed isotopic fractionations predicted by the Wing & Halevy model of microbial sulfate reduction (MSR). The Wing & Halevy model presumes that the net $^{34}\epsilon_{SO4/HS^-}$ from MSR is controlled by the reversibility of sulfate diffusion

into the cell and the three reductive steps of MSR (Figure 3.4). In all twenty simulations of MSR, QIRN reproduced the nonlinear behavior of $^{34}\epsilon_{SO4/HS^-}$ suppression as cSSRr increased, suggesting that QIRN's approach to isotopologue flux modelling accurately represents the balance between equilibrium and kinetic isotope effects.

Multiple isotope systems

Models of MSR also demonstrate the utility of QIRN for inorganic compounds and elements other than carbon. To model different elements, reactions will require multiple rows in the Reaction Database. For example, the transformation of alpha-aminonitrile to alanine has three carbon sites but only one nitrogen site to track. Each of these would be designated its own reaction. If the user calls the nitrogen-version of this reaction in their Network Builder, QIRN will automatically begin modelling its nitrogen isotope composition. As such, QIRN is generalizable to any isotope system involving two isotopes, including the $^{34}S/^{32}S$, $^{15}N/^{14}N$, $^{18}O/^{16}O$, and $^{2}H/^{1}H$ systems.

In simple networks, these isotope systems can be modelled simultaneously. For example, in MSR, both minor isotopes of sulfur (^{33}S and ^{34}S) could be modelled by assigning all sulfur species two atomic positions (despite sulfate only having one sulfur atom). One sulfur-site in the simulation would track $\delta^{34}S$ values and the other could track $\delta^{33}S$ values. Isotope effects would be expressed separately for the two isotope systems by assigning them to the first and second atomic positions in the Network Builder file. However, in QIRN's current version, a single atom cannot track more than two isotopes simultaneously because QIRN uses a binary representation of isotopologues (Equation 1). Thus, while there are strategies for modelling multiple isotope systems simultaneously, it is not yet a streamlined function of QIRN. This limitation may be overcome in future versions of QIRN by restructuring its isotopologue data storage methods.

Non-traditional isotopic properties

Modelling explicit isotopologues enables predictions of non-traditional isotopic properties that fall outside definitions of compound- or site-specific isotope compositions. For example, Orbitrap mass spectrometry can quantify the isotope ratios of molecular fragments (Neidhardt, 1987; Eiler et al., 2017). With the suite of isotopologues tracked by QIRN, modelling the isotope composition of these fragments is facile. While most of these properties have not yet been measured, the first step

in targeting fruitful non-traditional isotopic measurements is predictive modelling. QIRN provides the necessary computational power to make these predictions.

An example of non-traditional isotopic properties is intramolecular clumping anomalies from stochastic distributions. To our knowledge, these remain unexplored in larger molecules; however, they may become analytically accessible in the near future. When two carbon sites with different $\delta^{13}C$ are present in a molecule, the ^{13}C-^{13}C clumped composition of these sites will be depleted compared to the calculated stochastic abundance based on the average $\delta^{13}C$ of those sites, an artifact known as the combinatorial effect. Given the intense intramolecular $\delta^{13}C$ asymmetry in meteoritic alanine, a combinatorial ^{13}C-^{13}C effect is predicted between the C1 and C2 positions and between the C2 and C3 positions as compared to a stochastic distribution. After the Strecker synthesis network reached steady state, QIRN calculated a 6.5‰ ^{13}C-^{13}C clumping anomaly from a stochastic distribution between the C1 and C2 sites (Figure 3.3D). It also predictably changed with the isotope composition of the C2 site. The combinatorial effect could be corrected if the isotope compositions of the atomic sites were known. However, an independent measurement of the ^{13}C-^{13}C combinatorial effect on alanine's molecular fragments could be made via GC-Oribtrap by analyzing the average ^{13}R of the C1-C2 fragment and its clumped isotopologue ratio (Neubauer et al., 2018; Chimiak et al., 2021). This would serve as an independent constraint of alanine's intramolecular isotopic heterogeneity and synthetic origins without explicitly knowing its site-specific $\delta^{13}C$ values. Of course, this assumes nonstochastic KIEs on clumped isotopologues are not relevant, but clumped KIEs could also be modelled in QIRN (Equation 12). In this way, QIRN can predict isotopic properties of molecules that have never been measured, a useful exercise when directing the development of analytical techniques.

Case Studies

Photorespiration and the intramolecular ^{13}C patterns of plant glucose.

In simple networks, changes to branching ratios or the addition of new branches generally have predictable implications for the end-product isotope compositions. However, in cyclic networks, such as the Calvin Benson Bassham (CBB) cycle, small adjustments to fluxes can have cascading effects that produce non-intuitive signals in end products. With QIRN, users can now rapidly construct such networks and test variations as they explore their hypotheses. Here we used QIRN to explore *in*

silico the relationship between rate of photorespiration and glucose intramolecular ordering.

The CBB cycle drives the reduction of carbon dioxide into hexose sugars. Tcherkez et al. hypothesized that the relative fluxes and ^{13}C-fractionations of enzymes in this cycle were responsible for the non-stochastic intramolecular ^{13}C-distribution of plant glucose. They modelled the site-specific δ^{13}C values of glucose in the CBB cycle, treating δ^{13}C as an atomic property. The fluxes were independently constrained and the enzymatic isotope effects were optimized to fit previously measured site-specific glucose δ^{13}C values (measured in Rossmann, Butzenlechner, and H.-L. Schmidt, 1991) by Tcherkez et al. All of these were implemented into QIRN, including both enzymatic reactions and physical transport of metabolites between the chloroplast, peroxisome, and cytoplasm. As expected, the *in silico* predictions of glucose's intramolecular isotopic structure matched those of natural glucose (Figure 3.5). The only exception to this pattern was the C-6 site, which Tcherkez et al. removed from their optimization of enzymatic isotope effects. Thus, the isotopic composition of C-6 was not matched by either model.

Photorespiration (PR) is a side reaction of the enzyme Ribulose bisphosphate carboxylase/oxygenase (RuBisCo) in which RuBisCo reacts with O_2 instead of CO_2. PR sacrifices reduced carbon as CO_2 in the peroxisome, where the plant recovers three carbon sugars through glycine decarboxylase (GDC, Figure 3.5). Multiple previous experiments and modeling efforts have shown that as PR rises, plant glucose tends to become more ^{13}C-enriched, because GDC is highly fractionating and removes ^{12}C from the system as CO_2. While this result was apparent in QIRN simulations as well, QIRN's isotopologue approach enabled a more nuanced interpretation of glucose's site-specific δ^{13}C structure, analogous to the original Tcherkez et al. study.

In QIRN, higher rates of PR exaggerated the heterogeneities in glucose's δ^{13}C structure due to changes in the proportion of physical transport reactions between the chloroplast and cytoplasm. In particular, the C-4 position of glucose became almost 9‰ more enriched than the molecular average. Based on steady-state fluxes defined by Tcherkez et al. for the CBB cycle, fructose-6-phosphate (F6P) used for starch synthesis was exported from the chloroplast into the cytoplasm. This flux was not affected by increased PR. However, in response to higher loss of carbon via PR, the export of dihydroxyacetone (DHAP) decreased. Both exported DHAP and starch are sources of glucose in the cytoplasm. At high rates of PR, starch export becomes the primary source of F6P and glucose in the cytoplasm. This starch-derived F6P is

[13]C-enriched at the C-4 site due to the enzyme aldolase. Aldolase condenses G3P and DHAP into F6P and imposes a 16‰ inverse ^{13}C-fractionation on the C-1 position of G3P, as deduced by Tcherkez et al.. In QIRN simulations of 47% PR, the C-1 site of G3P in the chloroplast was 9.5‰ depleted from the molecular average, whereas cytoplasmic G3P exhibited intramolecular homogeneity. Aldolase's isotope effects were fully expressed in the chloroplast because it represents a small proportion of the branch point at the G3P node, unlike the analogous node in the cytoplasm, which has a 50/50 branching ratio. The excess ^{13}C lost from G3P was found in the product of aldolase: The C-4 site of starch-derived F6P and – subsequently glucose – in the cytoplasm. This enrichment occurred at all PR rates; however, as PR increased, more of the synthesized glucose was made from starch F6P and thus glucose's δ^{13}C structure became exaggerated.

The preceding discussion demonstrates the enormous interpretive power of isotopic models such as QIRN that can track site-specific isotopic compositions of all reactants in a complex network. Indeed, we are already poised to make predictions about how numerous isotopic properties in plants respond to growth rate, nutrient and water limitation, climate, and other parameters of interest (Wieloch et al., 2018). Many of these properties are only just beginning to be measured, if at all, and at present each new analyte and property requires significant research effort to develop appropriate protocols and standards. A more comprehensive model of plant physiology and biochemistry could help significantly to prioritize those analytical efforts to focus on the most useful isotopic signals.

Site-specific ^{13}C depletions in bacterial lipids

The ^{13}C depletion of lipids relative to other biomolecules is well documented, and is generally attributed to the ^{13}C kinetic isotope effect of pyruvate dehydrogenase (PDH). That enzyme decarboxylates pyruvate to acetyl-CoA, the building block molecule of most lipids (Melzer and O'Leary, 1987; Tang et al., 2017). From *in vitro* measurements of PDH's site-specific KIEs, it was predicted that fatty acids would inherit alternating ^{13}C enrichments and depletion at even and odd carbon numbers, respectively (DeNiro and Epstein, 1977). Monson and Hayes later confirmed this hypothesis by resolving intramolecular patterns of fatty acids from the bacterium *Escherichia coli*. Their models of *E. coli* grown on glucose were simple, involving only one branching point at pyruvate. Here we used QIRN to investigate whether their interpretations would hold in more detailed models of metabolism.

An independent model of *E. coli* carbon central metabolism, after Wijker et al., 2019, was constructed beginning with glucose as a substrate and including over 40 metabolic reactions. Enzymatic isotope effects led to an intramolecular pattern in acetyl-CoA of -14.1‰ and -10.5‰ on C-1 and C-2, respectively. Monson and Hayes deduced a ~6‰ deviation between theses sites. The discrepancy between modelled and measured results can be explained by the branching point at pyruvate, where it was estimated by Monson & Hayes that 74% of pyruvate flows through the PDH enzyme. Whereas in the Wijker et al. model, over 80% of pyruvate reacts with PDH, suppressing the enzymatic isotope effect further. However, the latter branching ratio proved to be a better fit for the carbon isotope composition of the C1 site in fatty acids. When the branching ratio at PDH was artificially decreased from 82% to 74% by increasing the flux from pyruvate to amino acids via acetolactate synthase, the RMSD between QIRN and *in vivo* C1 δ^{13}C values increased from 0.8‰ to 1.7‰. Thus, the 82% branching ratio determined by Wijker et al. was the better fit. Notably, this model of *E. coli* metabolism assumes all reactions are irreversible. Future studies could use QIRN to investigate the importance of enzymatic reversibility on downstream isotopic structures of central metabolites.

The extended *E. coli* model was then coupled to a fatty acid elongation pathway. It has been assumed that the intramolecular isotopic ordering of fatty acids would directly reflect that of acetyl-CoA. However, compound-specific and C-1 site-specific δ^{13}C values measured by Monson & Hayes supported downstream biosynthetic fractionations, which we have now modeled in detail with QIRN. Based on the δ^{13}C values for C-1 in C14:0 and C16:1 fatty acids, it was hypothesized by Monson & Hayes that a hydrolyzing enzyme was responsible for the isotopic differences between the individual fatty acids' C-1 positions. This enzyme stores fatty acids by diverting them from further elongation. A previous model deduced a 14‰ isotope effect on C-1 of fatty acids during its reaction. By assigning this fractionation on two different branching ratios for C14:0 and C16:1 fatty acids (7% and 75% hydrolyzed, respectively), Monson & Hayes explained the C-1 δ^{13}C depletion in both molecules. However, this model failed to capture how the residual ^{13}C-enrichments would impact downstream fatty acids (C16:0 and C18:1). In QIRN, the central metabolic model of *E. coli* was extended and the full lipid elongation pathway simulated. When the hydrolysis fluxes were assigned to match the relative abundance of *E. coli* fatty acids (Figure 3.6B), QIRN reproduced the the C-1 sites of all four major fatty acids. Had the 74% PDH branching ratio from the previous model been used instead, the C-1 position of C18:1 would have been >2‰ offset from the

measured value. Furthermore, the compound-specific values in QIRN simulations, which were not modelled in the prior study, have a pattern that matches *E. coli* cells. Using two independent models of *E. coli* metabolic pathways, the intramolecular ordering of its fatty acids were reconstructed, further validating the hypothesis by Monson & Hayes put forth 40 years ago. QIRN also adds a quantitative explanation for small differences in the compound-specific $\delta^{13}C$ of fatty acids, demonstrating its utility for interpreting highly complex networks.

The isotope distributions of bacterial anabolic products are influenced not only by the reaction network topology, but also by any intramolecular heterogeneity in supplied substrates (Bayle et al., 2015). In this study of *E. coli*, QIRN simulations assumed a homogeneous distribution of $\delta^{13}C$ in glucose to match the model put forth by Monson and Hayes. However, it is well documented that glucose does generally have a heterogeneous isotopic structure (Rossmann, Butzenlechner, and H.-L. Schmidt, 1991; Wieloch et al., 2018; Gilbert et al., 2012). To investigate this further, we introduced site-specific $\delta^{13}C$ variations in glucose in the QIRN model. For example, when the intramolecular $\delta^{13}C$ of glucose measured in sugar beet syrup (Figure 3.5, Rossmann, Butzenlechner, and H.-L. Schmidt, 1991) was used in the *E. coli* model, QIRN predicted that the C-2 site of acetyl-CoA would be 2.6‰ depleted compared to simulations with homogeneous glucose. As a result, the carbon sites of acetyl-CoA were only 1.9‰ offset from each other. Nevertheless, this particular simulation still matched *E. coli's* fatty acid molecular average and C-1 site-specific $\delta^{13}C$ values with an error of 1.2‰ and 0.8‰ (RMSD), respectively. When the ^{13}C-depleted C-6 site of glucose was instead set to the molecular average of sugar beet glucose, the intramolecular offset in acetyl-CoA increases to 4.3‰. Thus, it appears that glucose's intramolecular ^{13}C structure - particularly at C-6 - exerts strong control over resulting isotopic heterogeneities in fatty acids. Moreover, this influence cannot be discerned simply through measurements of the C-1 and molecular average $\delta^{13}C$ values. This result could help to explain why studies of yeast fatty acids did not find similar ^{13}C heterogeneities as in E. coli (Monson and J. M. Hayes, 1980). It also has meaningful implications for the astrobiology community, which points to the intramolecular carbon isotope ordering of long-chain lipids as a potential biosignature of life. However, it is clear from QIRN simulations that the fractionations Monson & Hayes observed could have been partially inherited from the glucose starting material. Future studies could use QIRN to reject the null hypothesis that isotopic properties of substrates are driving anomalies in anabolic products.

Limitations

QIRN's approach to modelling stable isotopes benefits greatly from its generalizability; however, this comes at the cost of four notable limitations: isotopologue array lengths, reaction network size, steady state requirements, and binary isotope formulations. These limitations are described for the current iteration of QIRN's programming.

Isotopologue Arrays

Tracking explicit isotopologues of a given binary isotope system requires that the program follow 2^n species for each molecule. Therefore, the lengths of arrays representing isotopologue populations increase exponentially with the size of the molecule. This becomes more cumbersome when molecular rearrangement functions, which loop through isotopologue arrays, are called. With the author's 16 GB unified memory laptop, QIRN failed to run when more than 10^6 isotopologues were tracked at once. Most molecules in central metabolism are <8 carbon atoms. Even for reaction networks with over 100 reactions, there are <30,000 isotopologues to track. Computer memory will only become limiting when larger molecules like lipids are modelled. However, with the support of high performance computing, longer isotopologue arrays will be feasible. Alternatively, users can avoid burdensome array lengths by approximating large molecules as smaller molecular fragments of interest. For example, the long-chain fatty acids in *E. coli* were modelled as C_4, C_6, and C_8 moieties by assuming that the first ten carbon atoms in each lipid were identical to five repeating acetyl-CoA molecules. Thus far, the largest molecule modelled in QIRN had 15 atomic sites and over 32,000 carbon isotopologues.

Network Size

Computational expense has a more obvious correlation with the number of reactions and intermediates in the network (Figure 3.7). The exact computational time for a single time step will be dependent on myriad external factors (computational power, available computer memory, etc.), but clearly increases as the reaction network expands. As noted above, the MSR network model ran significantly faster than the similarly-sized Strecker model, and we do not yet understand the reason(s) for this increased speed. One possibility is that the smaller MSR model, which only includes two isotopologues per substrate, has entered a regime where the memory handling of arrays becomes more efficient. We further note that – unlike all the
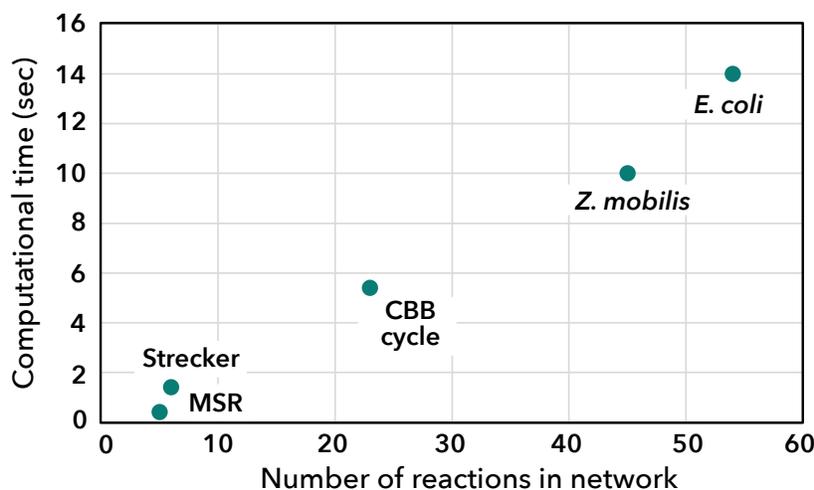
Figure 3.7: Computational time to run 10,000 time steps scales with the size of the reaction network. Abbreviations: MSR: microbial sulfate reduction, CBB: Calvin-Benson-Bassham

other networks – MSR has no intramolecular rearrangements, which may make it more efficient. Regardless, assuming a linear trend from Figure 3.7, a model of *E. coli*'s entire metabolism (~800 reactions) over 10,000 timesteps would require only ~560 seconds (Orth et al., 2011) using the author's laptop. This is trivially short for many applications, though long enough to become prohibitive for Monte Carlo type simulations with thousands of iterations. Even in the aforementioned model of MSR, where QIRN simulations run in less than 0.5 seconds, QIRN is still far more computationally expensive than solving the analytical solution for this network, which is functionally instantaneous.

Steady-State Extends Run Times

For many models of isotopic composition, attainment of isotopic steady state is an important assumption. In QIRN, each unique reaction network will require a different number of time steps to achieve that steady state. Importantly, QIRN always achieves a concentration steady state before – and sometimes long before – it achieves isotopic steady state. This is because the isotopic composition of substrates are dependent on reaction fluxes and branch points; the expressed isotope fractionation at each branch point will change until the fluxes and concentrations have stabilized. The time to reach isotopic steady state is dependent on a number of factors, but most importantly the turnover time of each substrate, defined as the reservoir size divided by the output flux. Molecules with longer turnover times take

longer to reach steady-state isotopic compositions (an asymptotic process) (John M Hayes, 2001). Therefore, a good strategy for decreasing run times is to keep the ratio of absolute flux-to-reservoir size of metabolites high, resulting in short turnover times. This also means that any products without reactions to consume them will continue to accumulate and never reach a steady-state composition. For products of interest, arbitrary sink reactions must be added, as was done for sulfide and alanine in the MSR and Strecker networks.

Binary Isotope Formulation

QIRN's method for tracking isotopologues employs a binary representation of isotopes. As a result, only two isotopes can be tracked at once. As described above (Section 4.2.2), there are strategies to circumvent this issue. However, they only apply to substrates with one atomic site. For larger molecules, a systematic approach to modelling three or more isotopes simultaneously is required. The binary isotope approach also limits modelled clumped isotopologues to those with multiple of the same rare isotope substitutions (i.e. $^{13}C - ^{13}C$ and not $^{13}C - ^{2}H$). By restructuring portions of QIRN's Python functions for tracking isotopologues, three or more stable isotopes could be represented for a single atom; however, this has not been implemented in the current version, in part because moving away from the binary representation will limit QIRN's inherent computational efficiency.

3.6   Conclusions

QIRN is a new modelling tool for interpreting and predicting natural-abundance stable isotope compositions in complex reaction networks. Here, five case studies demonstrated the applicability of QIRN to biological metabolisms and abiotic synthetic networks. In these examples, QIRN reproduced both previous models and data collected from natural samples. This confirms QIRN's ability to model highly complex networks that involve reversibility, clumped isotope anomalies, kinetic/equilibirium isotope effects, and a diverse set of physical, biological, and chemical processes.

QIRN enables stable isotope chemists to explore their applications with more quantitative power than before. Entire metabolisms and synthetic networks can be easily constructed without the need to create and solve large systems of equations, as was previously required. Once this network has been assigned by the user, QIRN builds it autonomously and reports many numerical solutions of isotopic properties simul-

taneously, lending it unprecedented flexibility. In the future, QIRN may also predict anomalous isotopic properties of molecules that correlate with important processes across many scientific disciplines.

Code Availability

Code to run QIRN and its GUI can be found in a public GitHub repository at: https://github.com/elliottmueller/QIRN-Files.git

Acknowledgements

References

Bayle, Kevin et al. (Feb. 2015). "Nonstatistical 13C Distribution during Carbon Transfer from Glucose to Ethanol during Fermentation Is Determined by the Catabolic Pathway Exploited". en. In: *Journal of Biological Chemistry* 290.7, pp. 4118–4128. DOI: `10.1074/jbc.M114.621441`.

Chimiak, L. et al. (Jan. 2021). "Carbon isotope evidence for the substrates and mechanisms of prebiotic synthesis in the early solar system". en. In: *Geochimica et Cosmochimica Acta* 292, pp. 188–202. DOI: `10.1016/j.gca.2020.09.026`.

Christensen, Bjarke and Jens Nielsen (Oct. 1999). "Isotopomer Analysis Using GC-MS". en. In: *Metabolic Engineering* 1.4, pp. 282–290. DOI: `10.1006/mben.1999.0117`.

DeNiro, Michael J. and Samuel Epstein (July 1977). "Mechanism of Carbon Isotope Fractionation Associated with Lipid Synthesis". en. In: *Science* 197.4300, pp. 261–263. DOI: `10.1126/science.327543`.

Douglas, Peter M.J. et al. (Nov. 2017). "Methane clumped isotopes: Progress and potential for a new isotopic tracer". en. In: *Organic Geochemistry* 113, pp. 262–282. DOI: `10.1016/j.orggeochem.2017.07.016`.

Eiler, John et al. (Nov. 2017). "Analysis of molecular isotopic structures at high precision and accuracy by Orbitrap mass spectrometry". en. In: *International Journal of Mass Spectrometry* 422, pp. 126–142. DOI: `10.1016/j.ijms.2017.10.002`.

Gilbert, Alexis (May 2021). "The Organic Isotopologue Frontier". en. In: *Annu. Rev. Earth Planet. Sci.* 49.1, pp. 435–464. DOI: `10.1146/annurev-earth-071420-053134`.

Gilbert, Alexis et al. (Oct. 2012). "Intramolecular ^{13}C pattern in hexoses from autotrophic and heterotrophic C $_3$ plant tissues". en. In: *Proc. Natl. Acad. Sci. U.S.A.* 109.44, pp. 18204–18209. DOI: 10.1073/pnas.1211149109.

Goldman, Mark Jacob et al. (June 2019). "Computer-generated isotope model achieves experimental accuracy of filiation for position-specific isotope analysis". en. In: *Chemical Geology* 514, pp. 1–9. DOI: 10.1016/j.chemgeo.2019.02.036.

Hayes, John M (2001). "Fractionation of the Isotopes of Carbon and Hydrogen in Biosynthetic Processes". en. In.

Jacobson, Tyler B. et al. (July 2019). "2H and 13C metabolic flux analysis elucidates in vivo thermodynamics of the ED pathway in Zymomonas mobilis". en. In: *Metabolic Engineering* 54, pp. 301–316. DOI: 10.1016/j.ymben.2019.05.006.

Jonathon Gropp Qusheng Jin, Itay Halevy (2022). "Controls on the isotopic composition of microbial methane". en. In: *SCIENCE ADVANCES*.

Melzer, Eva and Marion H. O'Leary (May 1987). "Anapleurotic CO $_2$ Fixation by Phosphoenolpyruvate Carboxylase in C $_3$ Plants". en. In: *Plant Physiol.* 84.1, pp. 58–60. DOI: 10.1104/pp.84.1.58.

Millard, Pierre, Jean-Charles Portais, and Pedro Mendes (Dec. 2015). "Impact of kinetic isotope effects in isotopic studies of metabolic systems". en. In: *BMC Syst Biol* 9.1, p. 64. DOI: 10.1186/s12918-015-0213-8.

Monson, K. D. and J. M. Hayes (1980). "Biosynthetic control of the natural abundance of carbon 13 at specific positions within fatty acids in Escherichia coli. Evidence regarding the coupling of fatty acid and phospholipid synthesis." In: *Journal of Biological Chemistry* 255.23, pp. 11435–11441. DOI: https://doi.org/10.1016/S0021-9258(19)70310-X.

Neidhardt, Frederic Carl, ed. (1987). *Escherichia coli and salmonella typhimurium: cellular and molecular biology. 2*. en. Vol. 2. Washington, D.C: American Society for Microbiology.

Neubauer, Cajetan et al. (Nov. 2018). "Scanning the isotopic structure of molecules by tandem mass spectrometry". en. In: *International Journal of Mass Spectrometry* 434, pp. 276–286. DOI: 10.1016/j.ijms.2018.08.001.

Nimmanwudipong, Tarin et al. (2016). "Determination of Intramolecular 13C Isotope Distribution of Pyruvate by Headspace Solid Phase Microextraction-Gas Chromatography-Pyrolysis-Gas Chromatography-Combustion- Isotope Ratio Mass Spectrometry (HS-SPMEGC- Py-GC-C-IRMS) Method". en. In: *J Anal Bioanal Tech* 7.1. DOI: 10.4172/2155-9872.1000293.

Ochs, Raymond S. and Tanaji T. Talele (Mar. 2020). "Revisiting prochirality". en. In: *Biochimie* 170, pp. 65–72. DOI: 10.1016/j.biochi.2019.12.009.

Ono, Shuhei, Jeemin H. Rhim, and Eric C. Ryberg (May 2022). "Rate limits and iso-topologue fractionations for microbial methanogenesis examined with combined pathway protein cost and isotopologue flow network models". en. In: *Geochimica et Cosmochimica Acta* 325, pp. 296–315. DOI: 10.1016/j.gca.2022.03.017.

Orth, Jeffrey D et al. (Jan. 2011). "A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism—2011". en. In: *Molecular Systems Biology* 7.1, p. 535. DOI: 10.1038/msb.2011.65.

Rhim, Jeemin H. and Shuhei Ono (Oct. 2022). "Combined carbon, hydrogen, and clumped isotope fractionations reveal differential reversibility of hydrogenotrophic methanogenesis in laboratory cultures". en. In: *Geochimica et Cosmochimica Acta* 335, pp. 383–399. DOI: 10.1016/j.gca.2022.07.027.

Rossmann, Andreas, Maria Butzenlechner, and Hanns-Ludwig Schmidt (June 1991). "Evidence for a Nonstatistical Carbon Isotope Distribution in Natural Glucose". en. In: *Plant Physiol.* 96.2, pp. 609–614. DOI: 10.1104/pp.96.2.609.

Schmidt, Karsten et al. (Sept. 1997). "Modeling isotopomer distributions in bio-chemical networks using isotopomer mapping matrices". en. In: *Biotechnol. Bioeng.* 55.6, pp. 831–840. DOI: 10.1002/(SICI)1097-0290(19970920)55:6<831::AID-BIT2>3.0.CO;2-H.

Sim, Min Sub et al. (Jan. 2019). "Role of APS reductase in biogeochemical sulfur isotope fractionation". en. In: *Nat Commun* 10.1, p. 44. DOI: 10.1038/s41467-018-07878-4.

Tang, T. et al. (Mar. 2017). "Geochemically distinct carbon isotope distributions in *Allochromatium vinosum* DSM 180 [T] grown photoautotrophically and photoheterotrophically". en. In: *Geobiology* 15.2, pp. 324–339. DOI: 10.1111/gbi.12221.

Tcherkez, Guillaume et al. (2004). "Theoretical considerations about carbon isotope distribution in glucose of C3 plants". en. In: *Functional Plant Biol.* 31.9, p. 857. DOI: 10.1071/FP04053.

Thomas, Burt, Katherine H. Freeman, and Michael A. Arthur (Feb. 2009). "In-tramolecular carbon isotopic analysis of acetic acid by direct injection of aqueous solution". en. In: *Organic Geochemistry* 40.2, pp. 195–200. DOI: 10.1016/j.orggeochem.2008.10.011.

Wang, David T. et al. (Apr. 2015). "Nonequilibrium clumped isotope signals in microbial methane". en. In: *Science* 348.6233, pp. 428–431. DOI: 10.1126/science.aaa4326.

Wasylenko, Thomas M. and Gregory Stephanopoulos (Sept. 2013). "Kinetic isotope effects significantly influence intracellular metabolite ^{13}C labeling patterns and flux determination". en. In: *Biotechnology Journal* 8.9, pp. 1080–1089. DOI: 10.1002/biot.201200276.

Wiechert, Wolfgang (2001). "13C Metabolic Flux Analysis". en. In: *Metabolic Engineering*.

Wieloch, Thomas et al. (Mar. 2018). "Intramolecular 13C analysis of tree rings provides multiple plant ecophysiology signals covering decades". en. In: *Sci Rep* 8.1, p. 5048. DOI: 10.1038/s41598-018-23422-2.

Wijker, Reto S. et al. (June 2019). "^{2}H/^{1}H variation in microbial lipids is controlled by NADPH metabolism". en. In: *Proc. Natl. Acad. Sci. U.S.A.* 116.25, pp. 12173–12182. DOI: 10.1073/pnas.1818372116.

Wilkes, Elise B. and Ann Pearson (Nov. 2019). "A general model for carbon isotopes in red-lineage phytoplankton: Interplay between unidirectional processes and fractionation by RubisCO". en. In: *Geochimica et Cosmochimica Acta* 265, pp. 163–181. DOI: 10.1016/j.gca.2019.08.043.

Wing, Boswell A. and Itay Halevy (Dec. 2014). "Intracellular metabolite levels shape sulfur isotope fractionation during microbial sulfate respiration". en. In: *Proc. Natl. Acad. Sci. U.S.A.* 111.51, pp. 18116–18125. DOI: 10.1073/pnas.1407502111.

Yoshida, Naohiro and Sakae Toyoda (May 2000). "Constraining the atmospheric N2O budget from intramolecular site preference in N2O isotopomers". en. In: *Nature* 405.6784, pp. 330–334. DOI: 10.1038/35012558.

C h a p t e r 4

METABOLIC CONTROLS ON CARBON ISOTOPE FRACTIONATIONS DURING BACTERIAL FERMENTATION.

Elliott P. Mueller[1*], Verena B. Heuer[2], Kai-Uwe Hinrichs[2], Alex L. Sessions[1]

[1]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA

[2]MARUM Centre for Marine Environmental Sciences, University of Bremen, Bremen, Germany
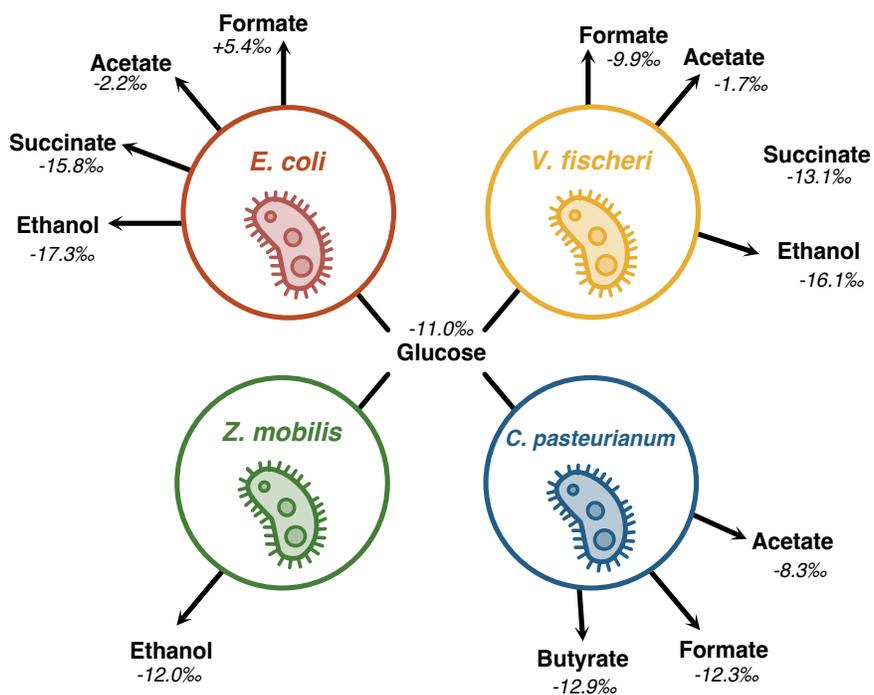
Graphical Abstract of Chapter 4

4.1 Abstract

The initial breakdown of organic matter into small molecules is thought to be the rate-limiting stage of anoxic organic degradation. Known as fermentation, this step is understudied in modern and ancient biogeochemistry due to a lack of environmental biomarkers (e.g. lipidomic, genomic) that would otherwise constrain its activity. It has long been assumed that fermentation, like respiration, does not express carbon isotope effects, precluding isotopic signals as a means of studying it in nature. Here, we tested this idea by growing pure cultures of four fermenting bacteria on glucose as a sole carbon source and measuring the carbon isotope composition of the organic acids and alcohols produced. We found that fermentation has a strong carbon isotope fractionation, ranging from -6‰ (normal) to +16‰ (inverse), depending on the fermentation product. This range can even be observed within a single organism. Using a combination of flux-balance analysis and bioisotopic models that track site-specific isotope enrichments through metabolic networks, we constrained the enzymes responsible for these fractionations. Our models reproduced the broad range of *in vivo* organic acid $\delta^{13}C$ values in all four organisms (model error = 1.6‰, n = 17). These findings demonstrate that acetate $\delta^{13}C$ enrichment is a consistent signature of fermentation that can explain environmental observations. Furthermore, the patterns of organic acid $\delta^{13}C$ values could serve as a fingerprint of the organic substrates being consumed *in situ* by fermenters. Looking to the geologic past, our results suggest that fermentation may have imprinted a trophic enrichment of ^{13}C isotopes on the three billion year record of sedimentary organic carbon, which may explain the inverse $\delta^{13}C$ pattern of acyclic isoprenoid and alkane biomarkers in Proterozoic sediments. Together, these results demonstrate the power of coupled metabolomics and bioisotopic modelling to understanding microbial isotope fractionations in modern and ancient biogeochemistry.

4.2 Introduction

The fate of organic matter in anoxic environments is a major lever on the Earth's carbon cycle. The mechanisms and rates of organic degradation underpin the Earth's long-term climate regulation (i.e. the burial of organic matter into ocean sediments) and short-term environmental responses to climate change (e.g. greenhouse gas emissions from wetlands) (DeVries, 2022; Reay et al., 2018). Fermentation is the first stage of anaerobic organic degradation during which biopolymers are broken down to small, labile molecules. Anaerobic microbial communities rely on these molecules for carbon and energy, which they could not otherwise access in the

form of primary produced organic matter (Arndt et al., 2013; Arnosti, 2004; Beulig et al., 2018). Fermentation products like acetate, formate, and ethanol have been observed in nearly every type of anoxic setting, highlighting the importance of this metabolism for carbon cycling. Acetate alone is the carbon source for as much as half of sulfate reduction and methanogenesis in the seafloor and two-thirds of methane production in terrestrial wetlands (Jørgensen, Findlay, and Pellerin, 2019; Conrad, 1999). Fermentation is a major driver of the global carbon cycle.

There are several important differences between fermentation and respiration that underlie the structure of anaerobic carbon cycles. Respiring metabolisms have an electron donor and an exogenous electron acceptor. As electrons are passed between these substrates, chemical potential energy is generated in the form of a membrane proton gradient, which can be dissipated to make ATP for the cell. However, fermentation does not typically rely on exogenous electron acceptors. Instead, it disproportionates organic carbon. For example, in the fermentation of glucose, a fraction of carbon is oxidized to CO_2 and another fraction is used as an electron acceptor, creating molecular hydrogen, organic acids and alcohols that are excreted into the extracellular environment. Contrary to respiring organisms, fermenting bacteria do not use electron transfer as their primary mode of energy conservation. Their largest energy source is substrate-level phosphorylation in central metabolism, which generates ATP through the cleavage of phosphorylated compounds(White, Drummond, and Fuqua, 2012). While this metabolism is less efficient in terms of ATP generation per mole of substrate, it seems to enable fermenting organisms to break down larger molecules. The carbon substrates that fermenting bacteria can utilize are diverse, including amino acids, sugars, and biopolymers (White, Drummond, and Fuqua, 2012). In contrast, at least in culture, anaerobically respiring bacteria and methanogens can only catabolize a small range of labile molecules. Iron reducing bacteria have been shown to break down sugars but they do not conserve energy in the iron reduction step. Rather, they do fermentation while exporting a small fraction of reducing equivalents to ferric iron (Derek R. Lovley and Phillips, 1986). The only substrates that are completely oxidized during iron reduction are fermentation products and select monoaromatic compounds (D R Lovley, 1991). These physiological constraints on metabolism underpin the ecological structure of anaerobic microbial food webs in which fermenters provide carbon and energy from organic matter to the rest of the community.

Fermentation has a strong influence on carbon biogeochemistry. Most importantly, it

may control the total rate of organic degradation and thus the rate of CO_2 and methane generation in a given environment (Beulig et al., 2018). Circumstantial evidence of this first appeared in the consistently low (μM) abundance of fermentation products like acetate across different settings. Organic acid generation and consumption rates were balanced, indicating that hydrolysis and fermentation were the rate-limiting stage of carbon degradation. Direct measurements of organic degradation rates throughout marine sediment profiles later confirmed this hypothesis (Arndt et al., 2013; Rothman, 2024). In one study, the total rate of organic oxidation across the sulfur-methane transition zone (SMTZ) was unaffected by the shift in terminal consumption process from sulfate reduction to methanogenesis (Beulig et al., 2018). Instead, the initial stages of organic matter decomposition set the pace of degradation throughout the sediment profiles. Fermentation slowly released organic acids that were rapidly — and competitively — converted to CO_2 and methane. These insights clearly demonstrate that hydrolysis and fermentation have the final word on carbon burial and methane production, increasing the demand for tools that can study these processes.

Despite its recognized importance, our ability to quantify fermentation in nature is limited. Open questions exist regarding the carbon substrates it uses, its metabolic activity and the microbial species performing fermentation. Oxidative substrates like sulfate and nitrate are routinely quantified in the environment as a means of tracing the metabolic activity of anaerobic respiration, but the organic substrates of fermentation are diverse and difficult to measure. The constant turnover of fermentation products makes their abundances equally difficult to interpret. To our knowledge, no culture-independent, biological markers of fermentation have been developed either. Fermenting organisms have numerous pathways to disproportionate carbon and these pathways are phylogenetically widespread (Hackmann and B. Zhang, 2023). As such, there are no characteristic enzymes or genes to screen for in environmental data. Metagenomic studies can only identify taxa that have been shown to perform fermentation in culture. More recent work to classify fermenters based on genotype have had some success, but they are also biased toward cultured, fast-growing gut microbiota (Hackmann and B. Zhang, 2023). Fermentation remains largely invisible to geochemical and biological tools in the environment.

Stable isotope fractionations between and within molecules have long been used to identify and even quantify metabolic activity. Indeed, the carbon isotope composition of organic acids have been characterized in many anoxic environments and

have demonstrably useful information. However, fermentation is assumed to have negligible isotopic fractionations during the production of these molecules (<3‰) (V. B. Heuer, Krüger, et al., 2010; V. Heuer et al., 2006; V. B. Heuer, Pohlman, et al., 2009; Penning and Ralf Conrad, 2006; Ralf Conrad, Liu, and Claus, 2021; Ralf Conrad and Claus, 2023). Their natural isotopic variability has instead been ascribed to consumption reactions, exchange with inorganic carbon or mixing with metabolic sources like autotrophic acetogenesis. (Lever et al., 2010; Franks et al., 2001)

Here, we sought to address two fundamental questions about fermentation within the framework of carbon isotope biogeochemistry: 1.) Does fermentation impose predictable isotopic fingerprint onto its products? 2.) Might these fingerprints be used to study fermentation in modern environments and in the rock record?

To answer these questions, we studied four model bacteria with distinct fermentation pathways. These organisms expressed large and variable isotope fractionations between glucose and the fermentation products (+16‰ to -6‰). We deduced the central metabolic fluxes of each organism with flux-balance analysis and implemented these networks into quantitative bioisotopic models specific to each organism constructed using Quantifying Isotopologue Reaction Networks (QIRN) (Mueller, Wu, and Sessions, 2022). Our models simulated the site-specific carbon isotope compositions of molecules within each organism's metabolic network. By comparing model outputs to *in vivo* data, we constrained the specific enzymatic reactions responsible for the observed fractionations, revealing the metabolic controls on the carbon isotope fractionations of fermentation. Understanding these controls enabled us to predict the range of possible isotopic fractionations given the constraints on fermentative metabolisms. These findings demonstrate that the pattern of organic acid $\delta^{13}C$ values could help trace the molecular identity of *in situ* fermentation substrates. Finally, we discuss the implications of a fermentation isotope effect on the carbon isotope composition of preserved organic matter in the Earth's rock record.

4.3 Methods
Culturing Conditions

Escherichia coli (MG1655), *Clostridium pasteurianum*, *Zymomonas mobilis*, and *Vibrio fischeri* (MJ11) were grown on minimal, defined media with glucose as a sole carbon source. Other than *V. fischeri*, phosphate buffer was used to maintain a pH of 7. For *V. fischeri*, MOPS was used as a buffer at pH 7.2. All media

compositions can be found in Tables 4.5-4.8. Cultures were initially spread on agar plates from glycerol stocks (frozen at -80C) in small, anoxic chambers. Since all organisms are facultative or spore-forming, plates were removed from anoxia to streak single colonies. Colonies were inoculated into rich-media containing 5 g/L yeast extract and the media components listed in Table 4.7 - 4.10 using balch tubes sealed with a butyl rubber stopper and tin crimp foil. These were sparged with nitrogen through a needle for 5 minutes. After growth, these were transferred to pre-flushed Balch-type tubes containing 5 mL of minimal media and glucose. These pure cultures were passaged 2-3 times before inoculation into triplicate, 1L culture bottle containing 100mL of minimal media and glucose (1-2% transfer volume) that had been flushed with nitrogen for 20 minutes. The glucose fed to these cultures was previously characterized for its site-specific $\delta^{13}C$ values using methods from Gilbert et al., 2009. Throughout the growth, optical density (OD) measurements at 600 nm were made by extracting 1 mL of sample through a syringe into a plastic cuvette and rapidly measuring the density on a spectrophotometer. Culture purity was confirmed via microscopy.

Organic acid quantification

At each growth timepoint, 0.5 mL of sample was removed, syringe filtered with a 0.22 μm filter and stored at -20C. These samples were diluted 2-fold into 8mN sulfuric acid and injected (10 μL) onto a high-performance liquid chromatography (HPLC) with a refractive index detector (RID) equipped with an autosampler. Sugars, organic acids and alcohols were separated on a Aminex-87H column with an isocratic method using 8mN sulfuric acid at 0.6 mL/min. The concentrations of glucose, succinate, acetate, ethanol, formate, and lactate were simultaneous quantified by running external standard curves (0.1-50 mM) of these analytes. In the case of *V. fischeri*, ethanol and MOPS co-eluted, precluding HPLC quantification of ethanol in those samples. To circumvent this issue, we measured ethanol on a gas chromatography (GC) instrument with flame ionization detection (FID). Here, samples were diluted two-fold into high-purity Milli-Q water and injected (0.3 μL) onto a ZB-WaxPlus column with a constant 120°C injector temperature (split ratio = 10). Concentrations were quantified with an external standard curve. Three water rinses of the autosampler syringe before and after injections ensured minimal sample carry over.

Organic acid isotopic analyses

Compound-specific isotopic analyses of sugars, volatile fatty acids and alcohols were performed on an isotope ratio monitoring liquid chromatography/mass spectrometry system (irm-LC/MS) at the University of Bremen. The analysis involved separation of compounds by high performance liquid chromatography (Thermo Scientific™ Dionex UltiMate 3000 HPLC) combined with chemical oxidation of the effluent using the Thermo Scientific™ LC IsoLink™ interface and subsequent online transfer of the resulting CO_2 into an irm-MS (ThermoFinnigan Delta Plus XP). The method was isocratic and isothermal. All solutions were aqueous, freshly prepared and degassed under vacuum in an ultrasonic bath (15 min at 40°C) in order to remove CO_2. The HPLC system was equipped with a VA 300/7.8 Nucleogel Sugar 810 H column (300 mm length; 7.8 mm i.d.) and a guard column (CC30/4 Nucleogel Sugar 810H; 30 mm length) from Macherey-Nagel, which were kept at 45°C. As a mobile phase, 5 mM phosphoric acid (prepared from 270 μL of 85% H_3PO_4 in 1 L of Milli-Q water) was used with a flow rate of 300 μL min^{-1}. The oxidation reactor temperature of the LC IsoLink™ interface was set to 99.9°C. Inside the interface, the eluent of the HPLC was mixed with oxidation reagent prepared from two solutions, one being 283 mM sodium peroxodisulfate solution (prepared from 27 g $Na_2S_2O_8$ in 400 mL of Milli-Q water) and the other 1.4 M phosphoric acid (prepared from 30 ml of 85% H_3PO_4 in 400 ml of Milli-Q water). The reagents were pumped into the interface with equal flow rates of 50 μL min^{-1}, adjusted to generate an oxidative power that yielded 11.5 V for oxygen measured as m/z 32 on cup number 2 (resistor of $3 \cdot 10^8$ ω). Isotope-ratio-monitoring was conducted on a ThermoFinnigan Delta Plus XP, to which the interface was connected with helium as carrier gas. Samples were injected using a temperature-controlled autosampler (set to 5°C), and a sample injection volume of 100 μl. Analysis were repeated with diluted samples when peak areas of one or several target compounds exceeded 500 Vs. In house isotopic standards of known $\delta^{13}C$ value were run at the beginning of the analytical sequence and in between every seven samples to ensure accuracy throughout. Precision ranged from 0.2‰ for ethanol, glucose and succinate to 0.7‰ for lactate and propionate. In samples, isotopic analysis of lactate suffered from co-elution with succinate when the latter was present in high concentrations. $\delta^{13}C$ values are reported by comparison to VPDB.

Lipid analyses

Fatty acid carbon isotope compositions were determined via gas chromatography isotope ratio mass spectrometer method. Fatty acids were first transesterified to fatty acid methyl esters (FAMEs). To do so, 1mL of GC-purity hexanes and 2 mL of a 20:1 mixture of anhydrous methanol and acetyl chloride were added to 5-10 mg of freeze-dried and crushed bacterial biomass. This mixture was heated to 100°C for 10 minutes. After allowing the solution to cool to room temperature, a liquid-liquid extraction with hexanes and water was performed. The organic phase was extracted and dried using a sodium sulfate column.

The FAME profile of each organism was determined using a Thermo Scientific Trace 1300 ISQ with separation on a 30-m × 0.25-mm capillary column (ZB-5 ms, 1 μm film thickness; Zebron). The $\delta^{13}C$ values of FAMEs were measured by a gas chromatograph coupled to an isotope-ratio mass spectrometer (Thermo-Scientific Delta+XP) using a combustion interface. Chromatographic separation was carried out with the same column. Identical oven ramp settings were used for GC/MS analysis, so that peaks could be identified on the GC-IRMS by relative retention time. For each sample, triplicate measurements were performed to determine $\delta^{13}C$ values. The average $\delta^{13}C$ values were reported relative to the Vienna Peedee Belemnite (VPDB) standard and were corrected for the added methyl C in the derivatization. An internal reference standard of CO_2 gas bracketed the chromatographic FAME peaks to determine their isotope composition on the VPDB scale.

Constructing flux-balance metabolic models

To construct the metabolic models for each organism, flux balance analysis was used. Previous investigations of fermentation pathways utilized by each bacterium were leveraged. From these studies, we created a set of four master flux maps to represent carbon flow in the organisms (Figure 4.1). The internal fluxes were constrained by the glucose uptake flux, fermentation product excretion fluxes and anabolic fluxes to key biosynthetic pathways including amino acids, fatty acids, and nucleic acids. We determined the later rates by assuming a cellular molecular composition (moles/cell) similar to that of *E. coli* as determined by Neidhartd (1982). By normalizing to the growth rate and cell densities, these were converted to flux terms (moles/g dry cells/hr). The excretion and uptake rates were constrained by plotting the concentration of these substrates and products versus the cell density for all biological replicates of each organism. The fitted linear slope of this plot was

taken as the excretion and uptake rates.

A system of equations to describe the internal fluxes of each cellular metabolic reactions was derived from the known stoichiometries of these reactions (Supplementary Information). Each system of equations for the different organisms is fully constrained by the empirical uptake, excretion, and anabolic fluxes (Supplementary Figure 4.8 and 4.10, Tables 4.2-4.4).

QIRN Models

A detailed description of QIRN's methods for modelling isotope fractionations within reaction networks is published elsewhere (Mueller, Wu, and Sessions, 2022). Briefly, we constructed the flux maps of fermentative metabolism that are abbreviated in Figure 4.1 and Figure 4.9. We assigned kinetic isotope effects to those enzymes listed in Table 4.1 at their designated atomic sites.

It was assumed that these KIE's all were the same in the four organisms. The $\delta^{13}C$ values of all the organisms' fermentation products could then be used as constraints on the KIEs. QIRN looped through a parameter space from 0.96 to 1 for each of these enzymatic KIEs, testing the residuals between model predictions and data for each simulation. These tests were performed in two-dimensions, changing two enzymatic KIEs independently and holding all others constant. The total parameter space was not searched as this would create intractably long computation times (months). All QIRN models were run for 2400 timesteps (dimensionless) which gave steady state isotope compositions for all end products. Rather than synthesizing a C_{18} fatty acid, which has over 250,000 isotopologues, we calculated its theoretical $\delta^{13}C$ value from the steady-state site-specific isotope compositions of its precursors, acetyl-CoA and malonlyl-CoA. Given the unidirectional nature of fatty acid biosynthesis, we assumed there was negligible fractionation from these monomers to the final lipid product. Palmitic acid is synthesized from an aceytl-CoA starting block and seven molecules of malonyl-CoA subsequently elongated off of this acetyl-CoA.

$$\delta^{13}C_{PA} = \frac{7 \times (\delta^{13}C_{MC,C1} + \delta^{13}C_{MC,C2}) + \delta^{13}C_{ACC,C1} + \delta^{13}C_{ACC,C2}}{16} \quad (4.1)$$
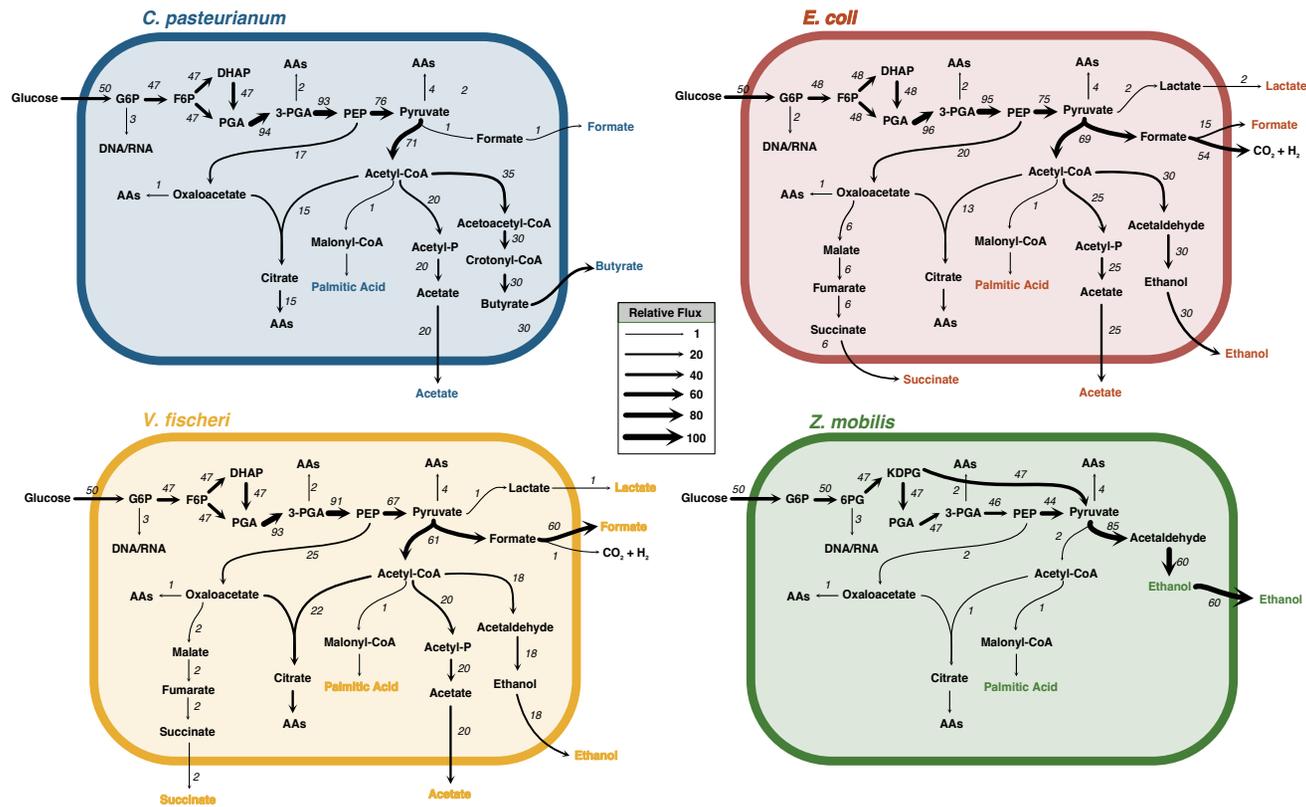
Figure 4.1: Metabolic flux maps of all four organisms grown and modelled using QIRN in this study. Unique branching ratios at each metabolic node explain their variable net isotopic fractionations without needing to alter enzymatic isotope effects between organisms. Certain reaction pathways involving multiple steps are represented by a single reaction in the diagram for brevity. Abbreviations: G6P; glucose-6-phosphate, F6P; fructose-6-phosphate, DHAP; dihydroxyacetone phosphate, 3-PGA; 3-phosphoglycerate, PEP; phosphoenolpyruvate, AA; amino acids.

Thus its $\delta^{13}C$ value($\delta^{13}C_{PA}$) is a function of the site-specific $\delta^{13}C$ values of malonyl-CoA's C1 and C2 sites ($\delta^{13}C_{MC,C1}$ and $\delta^{13}C_{MC,C2}$) and the site-specific $\delta^{13}C$ values of the C1 and C2 sites in acteyl-CoA ($\delta^{13}C_{ACC,C1}$ and $\delta^{13}C_{ACC,C2}$). We averaged these by 16, the number of carbons in palmitic acid, to derive the molecular average $\delta^{13}C$ value.

4.4   Results

Under fermenting conditions, all four organisms consumed glucose and excreted organics acids and alcohols. Growth rates for *E. coli*, *V. fischeri*, *C. pasteurianum* are similar at 0.47, 0.434, and 0.473 hr^{-1}. *Z. mobilis* had a slower (0.25 hr^{-1}) growth rate (Supplementary Figure 4.6). Two of the organisms (*E. coli* and *V. fischeri*) performed a mixed-acid fermentation, creating acetate, formate, ethanol, succinate, and lactate. *Z. mobilis* excreted exclusively ethanol, while *C. pasteurianum* made acetate, butyrate, and a minor amount of formate (Supplementary Figures 4.7 and 4.8). By tracking the time-varying concentrations of all the organic products, their excretion rates were calculated (mM/gcell/hr). Excretion rates were then normalized to the glucose uptake rate for each organism (mM/gcell/hr) to attain the zeroth-order fluxes found in Figure 4.1. These calculations were not possible for *E. coli* because time-varying samples were lost during transport. For *E. coli* , we assumed that the end-point substrate and product profile was representative of the uptake and excretion rates. From these data, a metabolic map of the metabolism was generated with flux-balance analysis (Figure 4.1). Under these assumptions and others regarding anabolism and reaction stoichiometry (see Methods), a unique metabolic network was determind for each organism studied. While similar in overall topology, the metabolic fluxes and branching ratios of each organism were different, even between *E. coli* and *V. fischeri* which have nominally the same mixed-acid fermentation pathway.

Each organism excreted organic products with different $\delta^{13}C$ values (Figure 4.2A). The glucose compound-specific $\delta^{13}C$ value (-11‰) did not deviate by >1‰ in any of the cultures during growth, suggesting that there were minimal isotope fractionations associated with substrate uptake (Table 4.2). The $\delta^{13}C$ value of acetate was consistently enriched compared to glucose by 2-9‰. Ethanol and succinate were ^{13}C-depleted (-16 to -18‰) in *E. coli* and *V. fischeri* , more so even than palmitic acid which was made by all four organisms. In *E. coli* , *V. fischeri* and *C. pasteurianum* , palmitic acid was 4-5‰ more ^{13}C-depleted than glucose. Meanwhile, ethanol and palmitic acid from *Z. mobilis* were both more enriched than in other organisms.

Formate created by *E. coli* had the largest isotopic fractionation from glucose with a +5‰ $\delta^{13}C$ value. Formate was also a major excretion pathway for *E. coli* . These results clearly demonstrated that fermentation expresses an isotope fractionation (Figure 4.2A). The fractionations were also consistent. The standard deviation $\delta^{13}C$ values from biological replicates were almost all less than 1‰, except for formate in *C. pasteurianum* (SD = 2.5‰). These fractionations did not arise from the uptake of glucose.
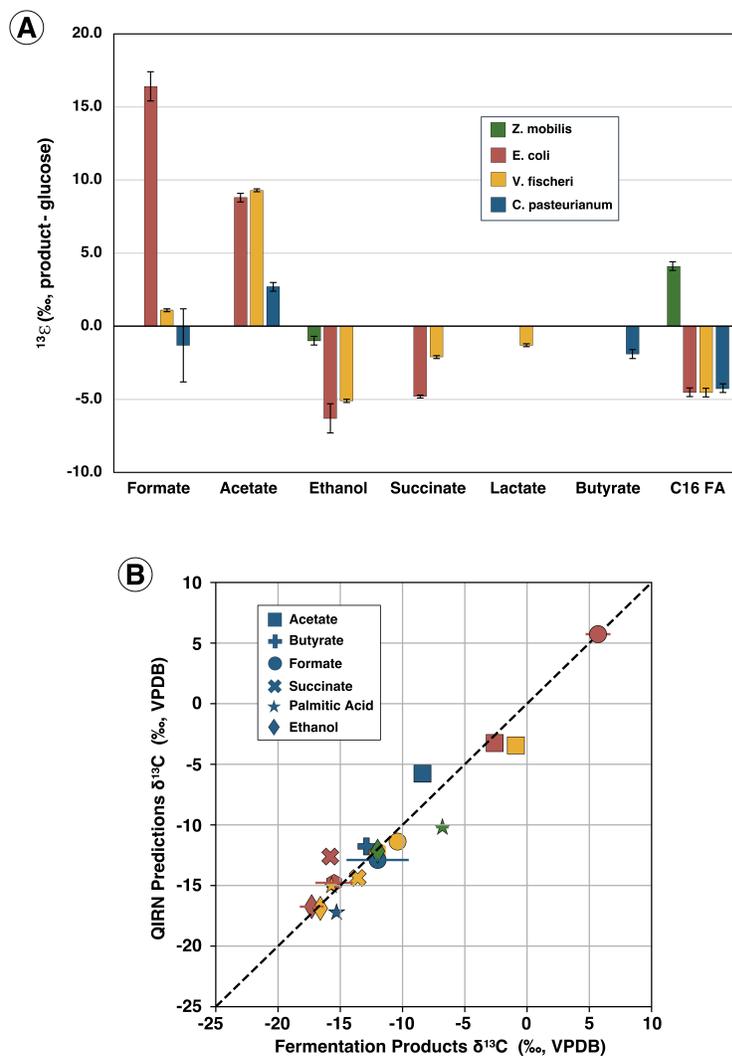
Figure 4.2: A.) Carbon isotope composition of fermentation products and palmitic acid (C16 FA) from the four fermenting organisms grown in this study. B.) Comparison of *in vivo* results and model predictions. Dotted line represents one-to-one line. Error bars on both plots represents standard deviation of biological triplicates.

Our next goal was to understand the enzymatic reactions that were controlling this

broad range of fractionations within and across organisms. To do this, we used a metabolic modeling tool called Quantifying Isotopologue Reaction Networks (QIRN). QIRN allowed us to model site-specific carbon isotope compositions of each metabolite in steady-state simulations of fermentation reaction networks (45-56 reactions and 48-60 metabolites). The models also incorporated the nonstochastic intramolecular $\delta^{13}C$ heterogeneity found within the glucose that these organisms consumed (Table 4.11). For more information on how QIRN constructs and simulates reaction networks, see Chapter 3. Model simulations of steady-state metabolic maps from Figure 4.1 were input into QIRN and site-specific ^{13}C kinetic isotope effects (KIE) were implemented into key enzymes within the fermentation pathways. With the fluxes determined, these KIEs became the free parameters of the models and were iteratively tested from 0.96 to 1.0 ($^{13}k/^{12}k$). For simplicity, secondary isotope effects were assumed to be negligible. The same enzymatic KIEs were applied to all four species. Model results were insensitive to KIEs for the majority of reactions (>80%), because they existed in unidirectional biosynthetic pathways (i.e. they had quantitative conversion of their substrate to product and could not express an isotope effect). However, the models were more sensitive to the KIEs of enzymes at the formate, pyruvate, and acetyl-CoA nodes of metabolism, like those of formate dehydrogenase (FDH), acetyl-CoA carboxylase (ACC), citrate synthase (CS), alcohol dehydrogenase (ADH), and phospho-acetyl transferase (PTA) (Table 4.1, Supplementary Figure 4.8). This was also a consequence of the data used to fit the model, which were pulled from the nodes of metabolism influenced by these enzymes. Other enzymes could be expressing isotope effects (e.g. to synthesize amino acids), but we do not have constraints on their KIEs so we do not model them here. The optimized KIEs found are in Table 4.1. Model outputs matched the data equally well within a range of 10‰ around the values listed in Table 4.1. QIRN fit the data (n = 17) with a root-mean square error (RMSE) of <1.7‰ (Figure 4.2B). This fit was possible even though we forced all four organisms to express the same enzymatic KIEs. The only difference in the models is the network topology. The ability to fit data across these four diverse organisms using the same set of isotope effects suggests that enzymatic KIEs between organisms are similar.

4.5   Discussion

Metabolic controls on $\delta^{13}C$ patterns of organic acids

The fate of acetyl-CoA controls carbon isotope fractionations during fermentation. Acetyl-Coa plays a central role in bacterial metabolism as an entry point to lipid

biosynthesis and the tricarboxylic acid (TCA) cycle. In fermenting bacteria, it is also the precursor for butyrate, ethanol and acetate production. Despite having a shared precursor molecule, all three of these products and the fatty acid palmitate had distinct $\delta^{13}C$ values. As such, the reactions consuming acetyl-CoA must have variable carbon isotope KIEs. In our models, the KIE of enzymes that initiate the TCA cycle, lipid biosynthesis, ethanol production and butyrate synthesis all have strong KIEs which leaves the acetyl-CoA pool ^{13}C-enriched (Table 4.1, Figure 4.3). Acetate inherits this ^{13}C-enriched signature, resulting in high $\delta^{13}C$ values of acetate in three of the bacteria. Meanwhile, fatty acids, ethanol and butyrate were ^{13}C-depleted in those same organisms. This inference would apply to fermentation pathways beyond those studied here, suggesting that ^{13}C-enrichment in acetate could be a consistent signal of fermentation.

In fermenting cells, the acetyl-CoA node balances ATP generation and redox control. Synthesizing acetate is a significant energetic boon for fermenters because it generates ATP. However, this pathway wastes acetyl-CoA molecules that could otherwise be used as an electron acceptor to produce ethanol or butyrate. These reactions regenerate NAD+ and other electron carriers that are required for glycolysis. Depending on the organism's genotype, its carbon substrate and its ambient environmental conditions, the branching ratio at acetyl-CoA will change. However, we predict that since the enzymes consuming acetyl-CoA have >10‰ isotope effects, the residual $\delta^{13}C$ of acetyl-CoA will always be enriched relative to the organic substrate (Figure 4.3). In one previous study of mixed-acid glucose fermentation, a similar acetate enrichment (4-5‰) was observed. Even when we grew *V. fischeri* and *E. coli* under respiring conditions, acetate was enriched relative to glucose, because citrate synthase was shuttling acetyl-CoA to the TCA cycle. Thus, regardless of how acetyl-CoA is partitioned, these reactions will express isotope effects that makes butyrate and ethanol ^{13}C-depleted and makes acetate ^{13}C-enriched (Figure 4.3A). As more carbon is excreted toward acetate and away from other acetyl-CoA pathways, acetate will approach the $\delta^{13}C$ value of the fermentation substrate but will not become more depleted. We contend that this will hold true when other carbon substrates (i.e. amino acids, polysaccharides) are fermented. Fermenting cells need to balance ATP generation, redox homeostasis and anabolism, all of which are controlled at the acetyl-CoA node. Thus, the physiology of fermenting prokaryotes predicts a consistent ^{13}C-enrichment of acetate relative to the consumed carbon substrates (Figure 4.3B).
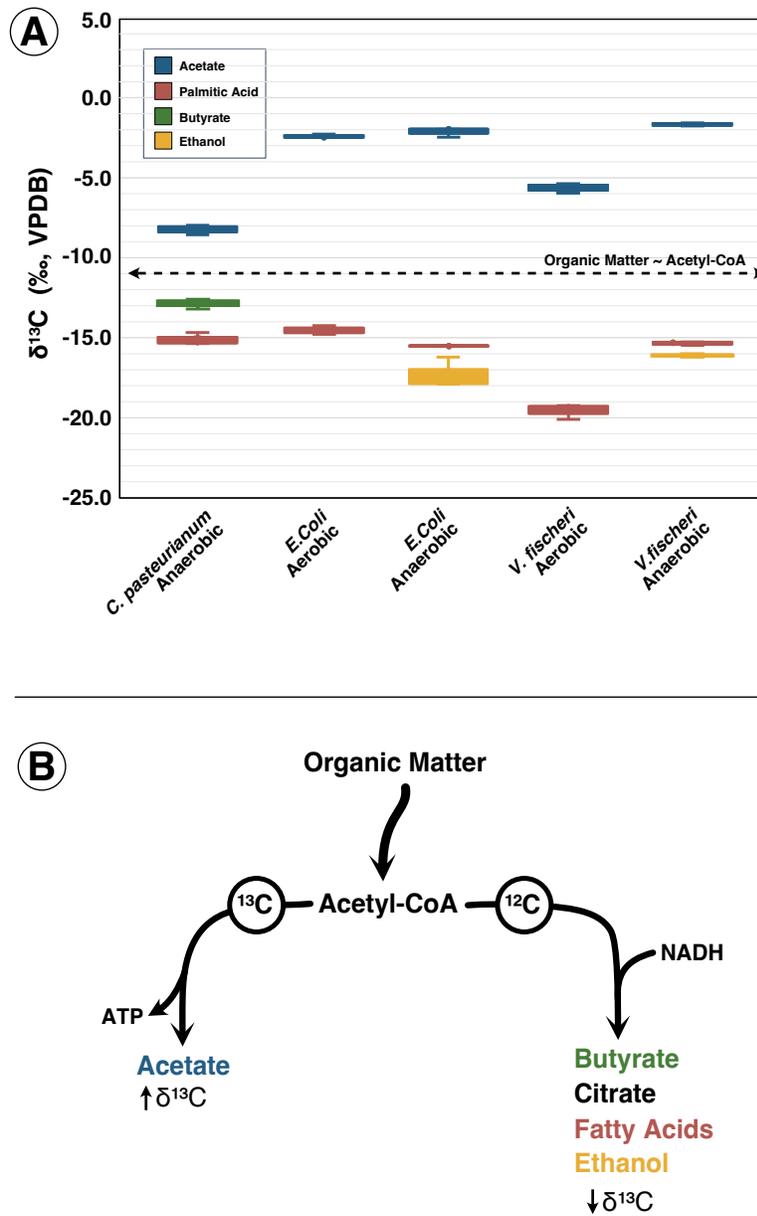
Figure 4.3: Comparison of carbon isotope composition in products formed from acetyl-CoA (top). Dotted line represents glucose molecular-average $\delta^{13}C$ composition. Box plots represent biological replicates. Schematic of carbon isotope flow at the acetyl-CoA node which leaves fatty acids ^{13}C-depleted and acetate ^{13}C-enriched. Carbon isotopes labels denote high magnitdue KIEs (<0.99)of enzymes to butyrate, citrate, fatty acids, and ethanol and low magnitude KIEs of enzymes to acetate (~1).

The preceding discussion warrants a reinterpretation of the carbon isotope signature found in bacterial acetogenic lipids, given that acetyl-CoA is also the building block of lipid biosynthesis. For decades, it has been assumed that fatty acids incorporate

Table 4.1: Best-fit KIEs for key enzymes in fermentative metabolisms.

Enzyme	Reaction	Site	KIE
ADH	acetyl-CoA $\rightarrow$ acetaldehyde	C-1	$0.973^{-0.007}_{+0.005}$
ACC	acetyl-CoA + CO_2 $\rightarrow$ malonyl-CoA	C-2	$0.973^{-0.005}_{+0.005}$
ACAT	2 x acetyl-CoA $\rightarrow$ acetoacetyl-CoA	C-2	$0.987^{-0.005}_{+0.005}$
		C-3	$0.987^{-0.005}_{+0.005}$
FDH	formate $\rightarrow$ CO_2	C-1	$0.976^{-0.005}_{+0.005}$
CS	oxaloacetate + acetyl-CoA $\rightarrow$ citrate	C-6	0.97-0.99
PTA	acetyl-CoA $\rightarrow$ acetate	C-1	1.000

the carbon isotope composition of acetyl-CoA with minor fractionations imposed during lipid elongation. Acetyl-CoA was thought to inherit ^{13}C-depleted carbon due to KIEs on the decarboxylation of pyruvate (Monson and J. M. Hayes, 1980; DeNiro and Epstein, 1977). However, our results demonstrate a consistent offset between acetate and palmitic acid $\delta^{13}C$ values (Figure 4.3). In bacteria, these molecules are generated from the same pool of acetyl-CoA. Thus, the reaction from acetyl-CoA to malonyl-CoA, known as acetyl-CoA carboxylase, must have a strong isotope effect. Otherwise, fatty acids would inherit the same ^{13}C-enriched carbon as acetate. The best fit across the four organisms is a C2 carbon isotope effect of 27‰. A KIE on acetyl-CoA carboxylase was similarly invoked to explain the intramolecular $\delta^{13}C$ pattern within plant-derived fatty acid; however, this model only required a 8‰ KIE (Julien et al., 2022). During *in vitro* studies of pyruvate carboxylase, which has an analogous reaction mechanism, the KIE on the methyl-site carbon isotope effect was 22‰, similar to our model results. We therefore propose a new hypothesis in which acetyl-CoA polymerization is responsible for the ^{13}C-depletion of fatty acid lipids in heterotrophic bacteria and possibly other organisms.

Interpreting acetate isotope compositions found in nature

Acetate is the best studied organic acid with respect to carbon isotope biogeochemistry. Over 95% of cultured fermenting bacteria synthesize it, and it appears in myriad anoxic environments (V. Heuer et al., 2006; Hackmann and B. Zhang, 2023). The high $\delta^{13}C$ values in acetate found in our study reflect observations from nature, including in oil brine fields, wetlands, rice-paddy fields and deep marine sediments (V. B. Heuer, Inagaki, et al., 2020; Ralf Conrad, Liu, and Claus, 2021). In the past, this enriched signature has been interpreted as either the incorporation of DIC into acetate's carboxylic acid group or an isotope effect imposed by con-

sumption metabolisms on the same carbon site. As a result, the methyl-site $\delta^{13}C$ is considered the most reliable signature of the precursor organic matter's carbon isotope composition. However, we demonstrate that a molecular-average inverse fractionation of 8‰ can derive from bacterial fermentation itself. Model outputs indicate that this fractionation is not localized to the C1 site, and the methyl-site may express fractionations as well (Figure 4.4). In fact, different intramolecular isotopic structures are predicted in acetate depending on the fermentation pathway that produced it (Figure 4.4). In concert with an understanding of which metabolisms are consuming acetate, its site-specific isotope composition may lend information about what fermentation pathways are active.

Constraining the substrates and pathways of fermentation

Fermentation provides more than one product that we can use to study its *in situ* activity. There is information stored in the carbon isotope compositions of its various products, a fingerprint of its metabolism. We specifically propose that the pattern of $\delta^{13}C$ values in fermentation products could identify the substrates utilized by fermenting bacteria in nature.

Organic matter has a complicated chemical composition (e.g. lipids, proteins, and sugars), which makes understanding its degradation pathways difficult. However, its different moieties have distinct $\delta^{13}C$ values, inherited from their initial biosynthesis (John M Hayes, 2001). Understanding the isotopic composition of the substrates being degraded could identify their molecular class or the locus of their formation (i.e. marine vs. terrestrial). Conservatively assuming that acetate excreted by fermentation has a $\delta^{13}C$ composition greater than or equal to its precursor, we can set a maximum $\delta^{13}C$ value for the organic matter being fermented. Given the branching at acetyl-CoA, we would also expect the substrate to have a $\delta^{13}C$ value higher than that of butyrate or ethanol. This creates a window of possible $\delta^{13}C$ values for the organic substrate. Co-occurring acetate and butyrate have been observed in nature, especially in permafrost soils where they accumulate to high concentrations (Drake et al., 2015). In tandem with models developed in this study, the $\delta^{13}C$ values of organic acids in permafrost could reveal the fermentation pathways that created them and the organic substrates that are being fermented *in situ*. However, these insights are currently limited by the small number of microorganisms tested in our study and the high KIE uncertainties. Thus, while we cannot fully unravel these signals yet, the methods used in this study represent a useful framework for moving
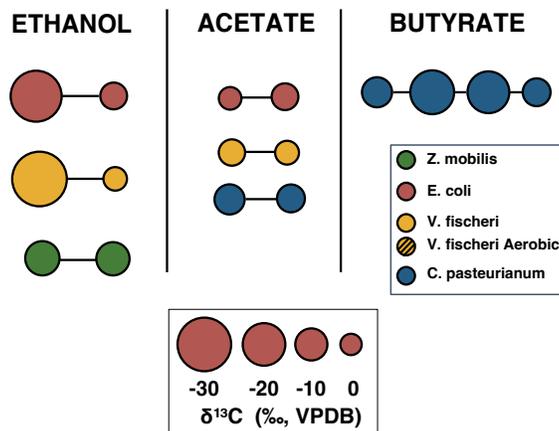
Figure 4.4: Prediction of the intramolecular isotopic structures within organic acids produced by fermentation.

this work forward.

Implications for ancient carbon biogeochemistry

The record of organic carbon burial within marine sedimentary rocks extends through the Archean era and is an important tool to study ancient life on Earth. These deposits have primarily been interpreted as a signal of autotrophy, with variable $\delta^{13}C$ values representing different proportions of carbon fixation pathways or changes to burial efficiency. In this study, we demonstrate that fermentative heterotrophy imposes its own isotopic fractionations which could be imprinted on the rock record, particularly in the Precambrian. In the following sections, we investigate potential signatures of this metabolism that are preserved in residual sedimentary organic matter.

Isotopic ordering of molecular fossils in the Precambrian

A common feature of ancient sedimentary deposits is the isotopic inversion of straight-chain and acyclic isoprenoid biomarkers across the Proterozoic-Phanerozoic boundary. This pattern is described as an "inversion" because it diverges from the *in vivo* isotopic ordering of biomolecules in algae and cyanobacteria. Modern phototrophs synthesize isoprenoids like phytol that are ^{13}C-enriched relative to fatty acid lipids (John M Hayes, 2001). These molecules are preserved as phytane and alkane biomarkers, respectively, and they retain their isotopic ordering ($\delta^{13}C_{alkanes}$ $<\delta^{13}C_{phytane}$) in Phanerozoic sediments. However, in sediments of the Proterozoic,

that ordering is consistently reversed ($\delta^{13}C_{\ alkanes} > \delta^{13}C_{\ phytane}$, Figure 5).

Since its discovery, several hypotheses have been generated to explain this isotopic inversion. They can be placed into two categories: A shift in the mechanism of primary production or heterotrophic reworking of organic matter. For example, Tang et al., 2017 demonstrated that by forcing a phototrophic bacterium to switch from photoautotrophic to photoheterotrophic growth (on acetate), they could recreate the Phanerozoic to Proterozoic isotopic inversion of fatty acids and phytol. However, they conclude that these signals are only likely to be preserved in the rock record as microbial mats where photoheterotrophy is active. Even this hypothesis implicitly invokes fermentation as a source of acetate for photoheterotrophy (Tang et al., 2017).

Alternatively, significant carbon degradation by anaerobic heterotrophic bacteria could explain the signals observed in Proterozoic sediments (Logan et al., 1995). Heterotrophs synthesize fatty acids for their cell membranes but do not produce phytol. If enough sedimentary carbon is converted into the biomass of secondary degraders (e.g. sulfate reducing bacteria), then the preserved alkane $\delta^{13}C$ signature would be a mixture of heterotrophic and autotrophic inputs, while the $\delta^{13}C$ value of phytane would record primary production alone. This hypothesis relies on an isotopic enrichment as carbon is transferred from primary producers to heterotrophs, known as a trophic enrichment. During aerobic respiration, trophic enrichments causes a net fractionation of 0 to 2‰, hence the commonly used term: "You are what you eat +1‰" (John M Hayes, 2001). However, based on models of sedimentary carbon degradation, heterotrophic-to-autotrophic ratios of lipids in sediments would have been too low to shift the $\delta^{13}C$ values of alkanes in Proterozoic sediments. In other words, muted enrichments from respiration are not enough to cause the observed isotopic inversion (Close, Bovee, and Pearson, 2011).

Trophic carbon isotope enrichments caused by fermentation

The major difference between Phanerozoic and Proterozoic deep water columns is oxygen concentrations. Ocean bottom waters experienced prolonged periods of anoxia throughout the Precambrian (Canfield, Poulton, and Narbonne, 2007). In those periods, anaerobic metabolisms would have had access to more labile carbon, and a significant fraction of primary produced biomass would have been degraded by fermentation rather than aerobic respiration. The inverse isotope fractionation caused by fermenting bacteria provides a mechanism for larger trophic ^{13}C enrichment that would only occur under anoxic conditions. If secondary de-
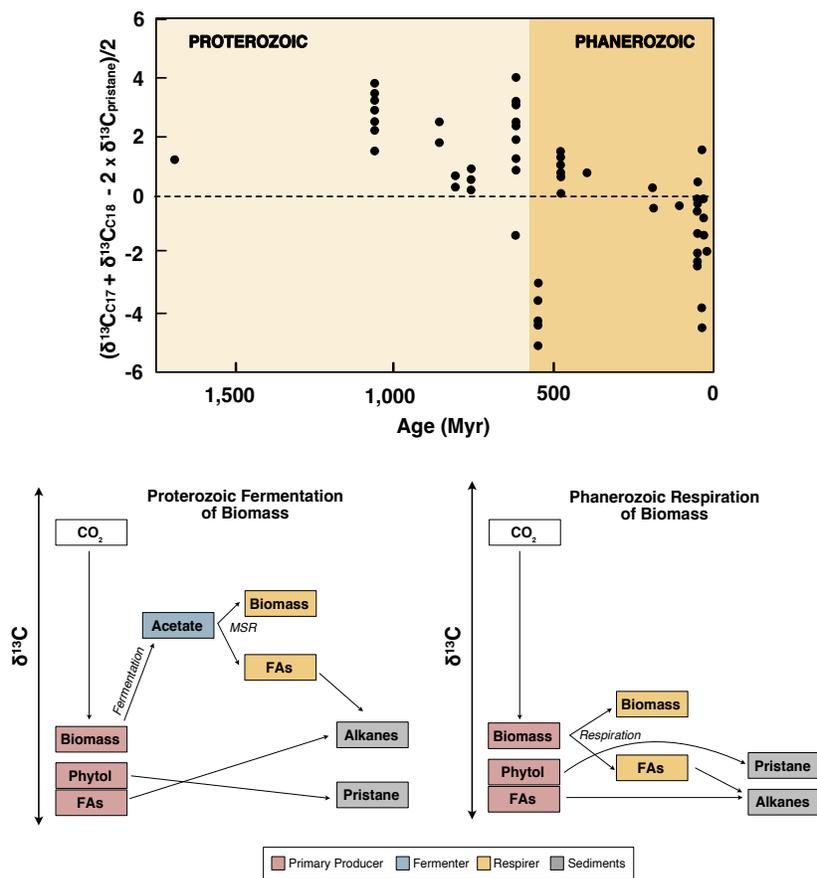
Figure 4.5: The isotopic inversion of acyclic isoprenoid biomarkers and n-alkanes across the Phanerozoic-Proterozoic boundary (top). A schematic model of organic carbon flow under anaerobic (bottom left) and aerobic (bottom right) conditions adapted from Logan et al., 1995. Inverse isotope fractionations imposed by fermentation would cause trophic enrichments of secondary degraders relative to primary producers that may explain the high $\delta^{13}C$ values of n-alkanes exclusively in the Proterozoic when deep ocean anoxia was widespread. All data from Logan et al., 1995 and references therein.

graders were consuming high $\delta^{13}C$ fermentation products, their lipids would inherit this signature. There are no conserved proteins unique to fermentative metabolism that can establish its presence in the Precambrian using molecular clocks. However, it is widespread on the tree of life, used by every domain of life and an estimated 30% of prokaryotes (Hackmann and B. Zhang, 2023). In fact, diverse cyanobacteria have been shown to ferment their own glycogen stores under dark, hypoxic conditions (Stal and Moezelaar, 1997; Klatt et al., 2020). Thus, in our model of trophic enrichments during the Proterozoic, strict heterotrophs or cyanobacteria — sinking below the photic zone or growing in regions of surface ocean anoxia — could equally have been the drivers of fermentation (Figure 4.5).

Trophic enrichments by fermentation could have been far larger than those imposed by aerobic respiration. Depending on the proportion of organic acids released from fermenting cells, the total $\delta^{13}C$ value of the organic acid pool would have changed, since more reduced products like butyrate and ethanol would likely have been ^{13}C-depleted while acetate would have been ^{13}C-enriched. Acetate is the most common microbial fermentation product and the most ubiquitous organic acid found in nature (Hackmann and B. Zhang, 2023). If acetate was the primary organic acid substrate for secondary degraders, the trophic enrichments could have been quite large, between 3 and 9‰. With more ^{13}C-enriched heterotrophic alkanes contributing to the sedimentary pool of lipids, the ratio of heterotrophic-to-autotrophic alkanes in sedimentary deposits required to yield the Proterozoic isotopic inversion would not have been as high as if aerobic respiration was the primary mode of carbon degradation. Recent studies of the 1.6 Ga Barney Creek Formation support this hypothesis. In those extracts, alkanes and C_{30} hopanes (a general biomarker of bacteria), had similarly enriched carbon isotope compositions compared to syngenetic phytane, providing further evidence that large trophic enrichments could be inherited by heterotrophic biomass during the Precambrian (X. Zhang et al., 2023). Conversely, in the Phanerozoic, fermentation would have been relegated to marine sediments where only the most recalcitrant organic compounds are deposited. In these deposits (as in modern sediments), only a small fraction of organic matter was processed through fermentation. Theoretically, a small portion of the alkane pool would still derive from ^{13}C-enriched secondary degraders. However, aerobic respiration served (and continues to serve) as the major heterotrophic metabolism, causing the trophic enrichment to disappear and solidifying the modern isotopic ordering of biomarkers.

Trophic enrichments specific to anoxic conditions have even been observed in mod-

ern setting (Gonzalez-Nayeck et al., 2023; Moran et al., 2023; Klatt et al., 2020). Gonzalez-Nayeck et al. 2023 studied a hypoxic microbial mat as an analogue of Proterozoic ecosystems and found that proteins from sulfate reducing bacteria living within the mat were ^{13}C-enriched relative to proteins in the primary producing cyanobacteria. In other studies, fermentation has been invoked to explain isotopic enrichments of lipids from green sulfur bacteria within anoxic mats (Van Bodegom, 2007). However, in oxic microbial mats, those trophic differences disappear. The $\delta^{13}C$ value of proteins from aerobic heterotrophs and cyanobacteria in the same mat were indistinguishable (Gonzalez-Nayeck et al., 2022). In summary, net organic carbon flow in oxic and anoxic conditions are similar (organics remineralized to CO_2), but the path of carbon flow (respiration vs. fermentation) has implications for the biomarker $\delta^{13}C$ record.

It may initially seem that quantitative conversion of organic material to fermentation products would negate the trophic enrichments described above. While this would be true for a unidirectional pathway (e.g. photosynthesis under low CO_2 concentrations), the isotopic fractionations of fermentation are imposed by metabolic branching required to maintain redox balance in the cell. In this case, the amount of total carbon consumed does not influence the overall carbon isotope fractionation. For example, each of the biological replicates of *C. pasteurianum* were captured at slightly different stages of exponential growth (due to variability in their individual lag times, Supplementary Figure 4.6), yet the $\delta^{13}C$ values of acetate and butyrate were consistent across the replicates (Figure 4.2). The net fractionations are determined by the ratio of organic acids, anabolic products, and CO_2 released, not the reaction progress. Layered onto metabolic fractionations are differences in organic substrate $\delta^{13}C$ values. Sugar monomers like glucose make up a large component of cyanobacterial biomass and are isotopically enriched relative to proteins or lipids (Van Bodegom, 2007; John M Hayes, 2001; Stal and Moezelaar, 1997). These sugars are easily degradable via fermentation and would likely have been the first compounds fermented. Thus, our hypothesis is only enhanced by selective degradation of labile carbohydrates.

In order to fully elucidate the prescence and magnitude of a fermentative trophic enrichment, degradation experiments need to be performed. Future work should focus on developing such model systems of anaerobic carbon degradation. With those insights, trophic enrichments identified by compound-specific $\delta^{13}C$ measurements of biomarkers could be a useful signature of fermentation in the rock record.

Conclusions

Fermentation plays a crucial role in the Earth' carbon cycle. Given its widespread importance across diverse ecosystems, tools to understand the activity and metabolic strategies of fermenting microorganisms are needed, both in modern environments and in the rock record. We demonstrate that fermenting bacteria growing on a simple sugar can express a large range of carbon isotope fractionations that are inherited by their excreted products. Using a combination of flux-balance analysis and bioisotopic modeling, we isolated the enzymes responsible for these fractionations. Given the physiology of fermenting bacteria, we predict which of these signals will be robust in the environment. We conclude that the patterns of $\delta^{13}C$ values in organic acids and alcohols could serve as a fingerprint of the substrates and metabolic pathways fermenters are employing to degrade organic matter. Looking to the geologic past, fermentation may have overlain a trophic enrichment of carbon isotopes onto the Earth's sedimentary rock record during periods of anoxia that can explain the isotopic ordering of biomarkers in the Proterozoic.

More broadly, our study demonstrates the utility of coupled metabolic models and isotopic measurements at their natural abundances. By identifying the mechanisms of isotope fractionations in microbial metabolism, we can evaluate the robustness of these signatures in the environment, generalize their application beyond pure cultures, and make predictions about how those signals will be preserved in the rock record.

4.6   Supplementary Materials

Growth was measured as optical density (OD) for each organism from lag phase to mid-exponential phase, at which point the cultures were rapidly collected for biomass. Growth rates were calculated from the exponential portion of each growth curve.
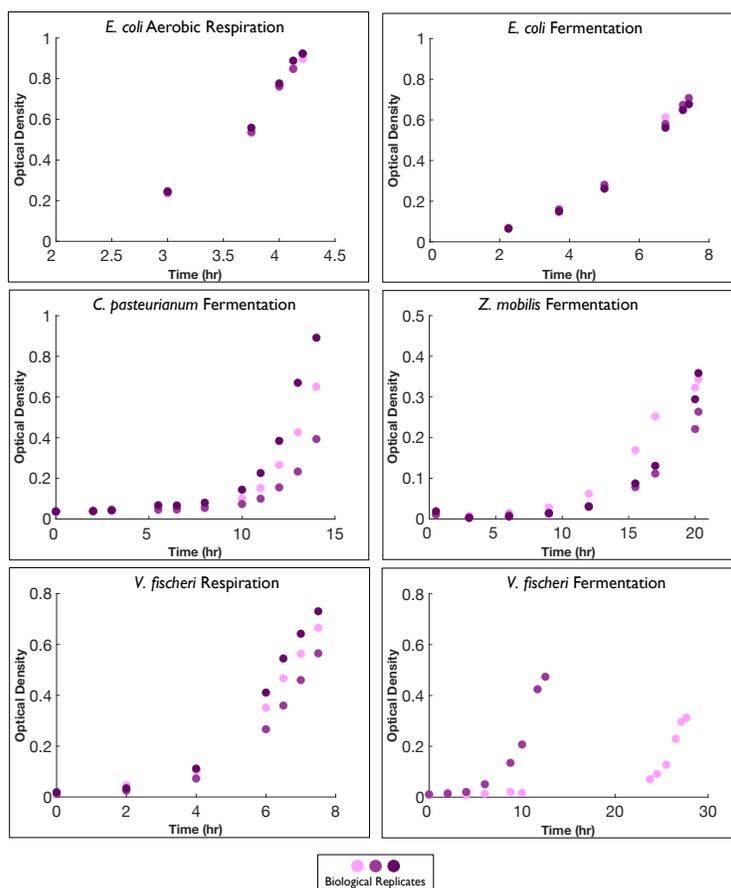
Figure 4.6: Growth curves of the four model organisms, *Escherichia coli*, *Vibrio fischeri*, *Zymomonas mobilis,* and *Clostridium pasteurianum*. *E. coli* and *V. fischeri* were grown aerobically and anaerobically. Colors indicate biological duplicate or triplicates.

Organic acid profiles were measured via high-performance liquid chromatography (HPLC) at the point of collection. Biological triplicates demonstrate high biological reproducibility.
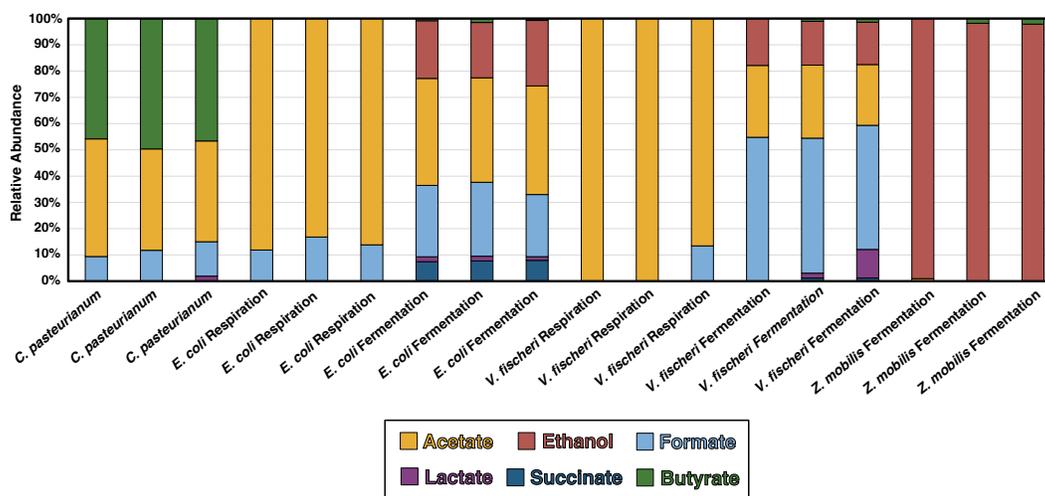
Figure 4.7: Relative abundance of fermentation product concentrations in mid-exponential phase at the point of biomass collection.
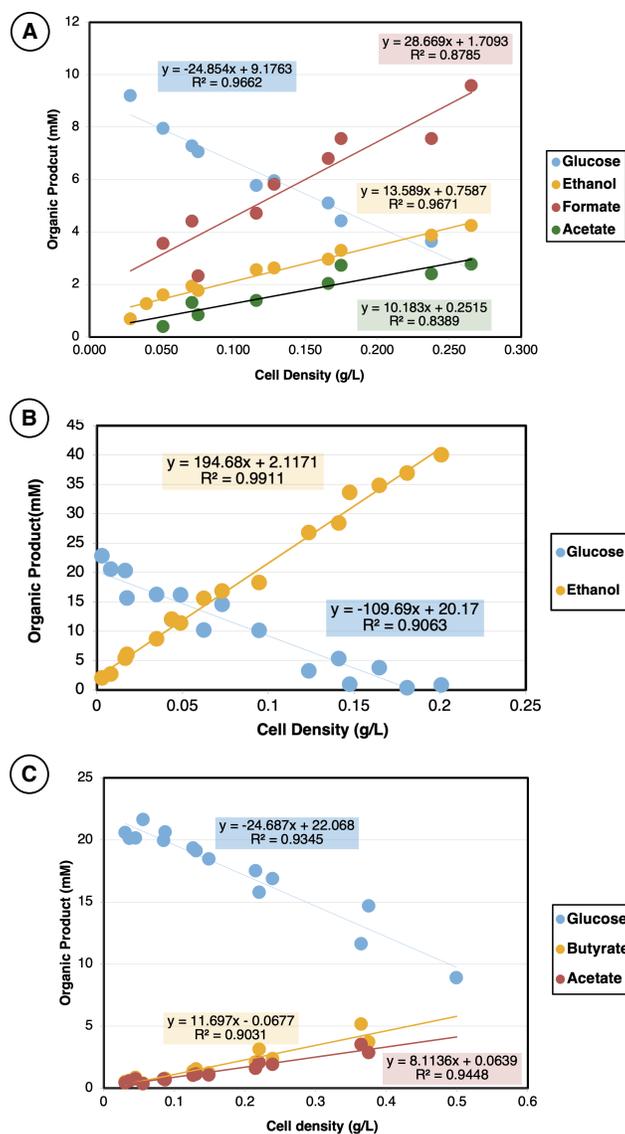
Figure 4.8: Relationships between time-varying cell densities (grams cell dry weight per liter), glucose concentration, and fermentation product concentrations. The slopes of these correlations are the zeroth-order fluxes used to constrain the flux balance analysis in terms of moles per gram cell dry weight per hour. A.) *V. fischeri* B.) *Z. mobilis* C.) *C. pasteurianum*

Fermentation products and glucose concentrations were measured throughout the growth to quantify excretion and uptake rates, respectively. Here, data from biological replicates were collated into a single data set to constrain the slopes in Figure 4.7.
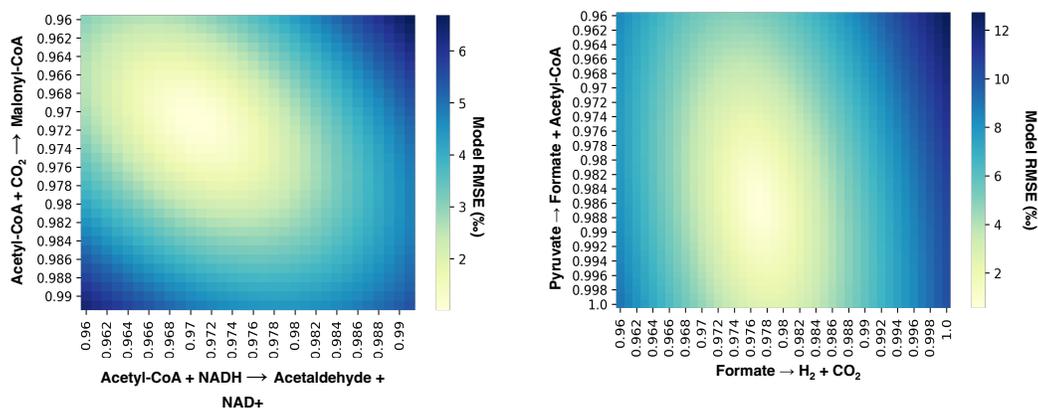
Figure 4.9: Sensitivity analysis of four enzymes kinetic isotope effects. Color gradient indicates the model-data residuals calculated as root-mean-square-error (RMSE). Carbon site of KIE is indicated in parentheses next to the enzyme name.

To evaluate the sensitivity of QIRN models to individual enzymes' KIEs, we varied two enzymatic KIEs independently and calculated the model-data residual as the root-mean square error with each simulation. In Figure 4.8A, carbon isotope data from fatty acids, acetate, ethanol and butyrate were used as constraints. In Figure 4.8B, formate and lactate carbon isotope compositions were used as constraints. To define the range of optimized KIE values (Table 4.1, in the main text), we set a limit of 3‰ for the RMSE which is three-fold higher than the analytical error of 1‰. These analyses demonstrated that acetyl-CoA carboxylase, alcohol dehydrogenase, and formate dehydrogenase can be constrained to varying extents. Acetyl-coenzyme A acetyltransferase is similarly constrained though not shown. Thus, while QIRN allows us to isolate the enzymes that must have kinetic isotope effects, it does not uniquely constrain the exact isotope fractionation on those enzymes. Nor does QIRN identify which carbon site(s) the KIE is occurring on. These isotope effects are placed exclusively at carbon sites that are involved in the reactions (e.g. bond rearrangements, hybridization changes, bond cleavage, or condensation).

Table 4.2: Compound-specific $\delta^{13}C$ values of excreted fermentation products from anaerobic cultures (‰, VPDB). EC; *E. coli* , VF; *V. fischeri* , ZM; *Z. mobilis* , CP; *C. pasteurianum* . A, B and C represent biological replicates.

Products	CP-A	CP-B	CP-C	EC-A	EC-B	EC-C	VF-A	VF-B	ZM-A	ZM-B	ZM-C
Glucose	-11.3	—	-10.7	-10.8	-10.5	—	-10.3	-10.6	—	-9.7	-9.6
Formate	-13.4	-9.4	-14.0	4.2	6.2	5.7	-9.8	-9.9	—	—	—
Acetate	-8.2	-8.6	-8.0	-2.5	-2.0	-2.0	-1.8	-1.6	—	—	—
Ethanol	—	—	—	-16.2	-17.9	-17.8	-16.2	-16.0	-12.0	-12.0	-11.9
Succinate											
Lactate	—	—	—	-3.0*	-5.3*	-5.8*	—	-12.3	—	—	—
Butyrate	-13.2	-12.6	-12.8	—	—	—	—	—	—	—	—
Paltimic Acid	-15.3	-14.7	-15.4	-15.5	-15.5	-15.5	-15.3	-15.5	-6.9	-7.0	-6.7

Table 4.3: Compound-specific $\delta^{13}C$ values of excreted acids from aerobic cultures (‰, VPDB). EC; *E. coli* , VF; *V. fischeri* . A, B and C represent biological replicates.

Products	EC-A	EC-B	EC-C	VF-A	VF-B	VF-C
Formate	-14.6	-14.9	-14.1	—	—	—
Acetate	-2.5	-2.5	-2.3	-5.6	-6.0	-5.4
Paltimic Acid	-14.3	-14.6	-14.8	-19.4	-19.2	-20.1

Table 4.4: Cell Composition: Amino Acids (Data from Neidhardt, 1987)

Amino Acids	μmol/g cell		μmol/g cell	μmol/g cell
Alanine	468	Lysine	326	
Arginine	281	Methionine	146	
Aspartic Acid	458	Phenylalanine	176	
Cysteine	87	Proline	210	
Glutamate	500	Serine	205	
Glycine	582	Threonine	241	
Histidine	90	Tryptophan	54	
Isoleucine	276	Tyrosine	131	
Leucine	428	Valine	402	

Amino Acid Family	Amino Acids
Pyruvate	ALA, VAL, ISO, LEU
GAP	GLY, SER
OAA	ASX, ARG
CIT	GLX, MET, PRO, THR
R5P	HIS, PHE, TRP, TYR

The anabolic fluxes used as constraints on the master metabolic maps of all four organisms were taken from (Neidhardt, 1987) and assumed to be consistent across all four organisms. The concentration of each anabolic product was normalized by growth rate and cell density to give a final zeroth-order flux of moles per gram cell dry weight per hour.

Table 4.5: Cell Composition: Nucleotides (Data from Neidhardt, 1987)

Nucleotides	μmol/g cell	Nucleotides	μmol/g cell
AMP	165	DAMP	24.6
GMP	203	DGMP	25.4
CMP	126	DCMP	25.4
UMP	136	DTMP	24.6

Table 4.6: Cell Composition: Lipids (Data from Neidhardt, 1987)

Lipids	μmol/g cell
Fatty Acids	258
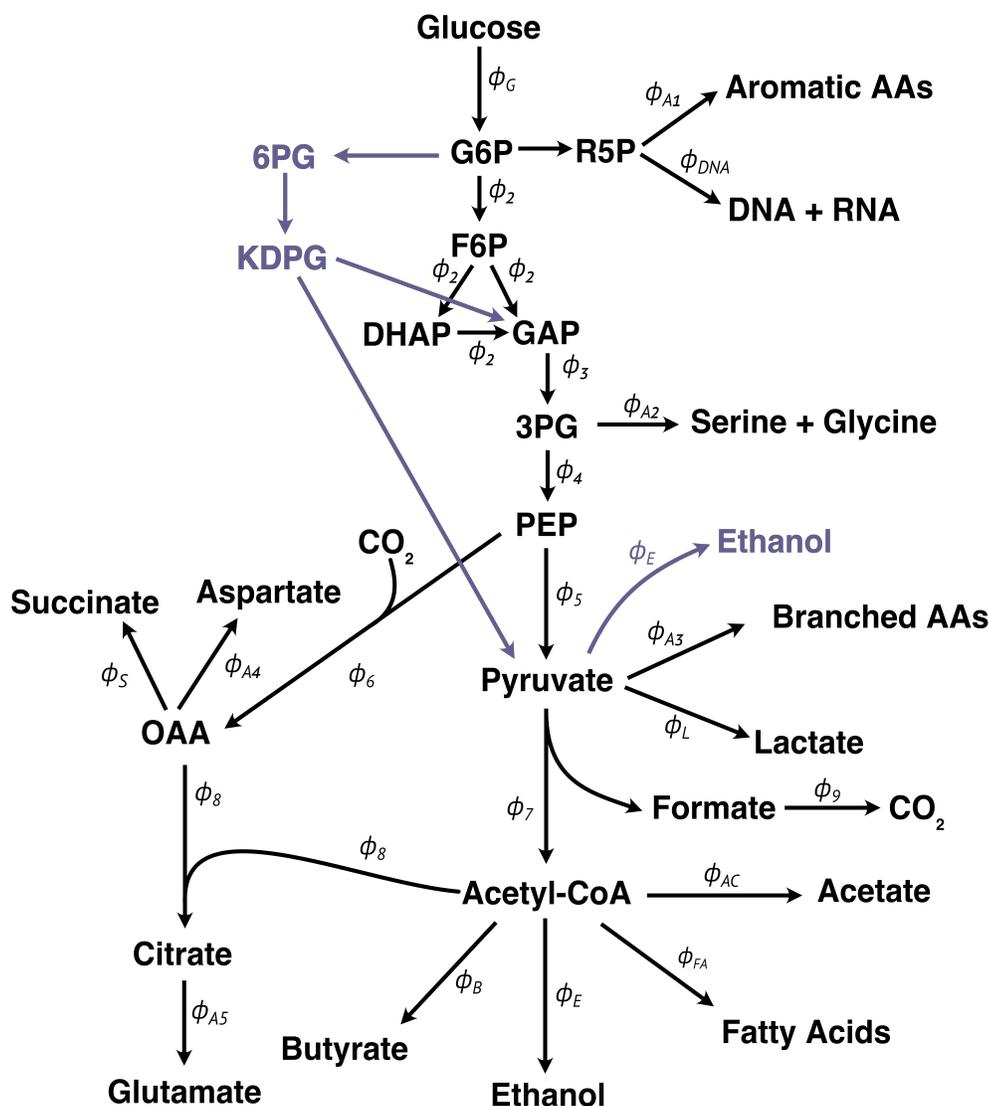
Figure 4.10: Master metabolic map for the four organisms grown in this study. Fluxes (ϕ) denoted with alphabetical subscripts represent anabolic and excretion fluxes that are estimated and empirically derived, respectively. Internal fluxes deduced from the system of equations below are denoted by numerical subscripts. Purple reactions denote those utilized by *Z. mobilis* but not the other organisms in this study.

Table 4.7: *Z. mobilis* Media

Component	Concentration (g/L)
$MgCl_2 \cdot 6H_2O$	0.4
NaCl	1
$CaCl_2 \cdot 6H_2O$	0.1
KCl	0.5
$CaCl_2 \cdot 2H_2O$	0.07
NH_4Cl	0.53
Na_2SO_4	0.04
K_2HPO_4	0.17
KH_2PO_4	0
Glucose	3.6
Ca-pantothenate	0.0953 (mg/L)

Table 4.8: *V. fischeri* Media

Component	Concentration (g/L)
$MgCl_2 \cdot 6H_2O$	8.73
NaCl	17.46
$CaCl_2 \cdot 6H_2O$	0.13
KCl	0.44
$CaCl_2 \cdot 2H_2O$	0.09
NH_4Cl	0.8
Na_2SO_4	3.55
K_2HPO_4	0.17
KH_2PO_4	0
Glucose	1.8

Escherichia coli, Vibrio fischeri, Zymomonas mobilis, and *Clostridium pasteurianum* were grown under fermenting conditions with glucose as a sole carbon source. The media conditions for each organism are below in Tables 4.2 to 4.5. All organisms were grown on a glucose stock with a known intramolecular isotopic structure (Table 4.11)

Table 4.9: *E. coli* Media

Component	Concentration (g/L)
Na_2HPO_4	11.32
KH_2PO_4	3
NaCl	0.5
NH_4Cl	2.5
$MgSO_4$	0.12
$CaCl_2$	0.011
Glucose	3.6

Table 4.10: *C. pasteurianum* Media

Component	Concentration (g/L)
Na_2HPO_4	2.2
KH_2PO_4	5.97
$(NH_4)_2SO_4$	6
$MgSO_4 \cdot 7H_2O$	0.12
$CaCl_2$	0.011
$Na_3Citrate \cdot 2H_2O$	0.33
Glucose	3.6

Table 4.11: Site-specific $\delta^{13}C$ values of glucose fed to all organisms.

Carbon Number	$\delta^{13}C$ (‰)
1	-6
2	-7.8
3	-8.5
4	-14
5	-14.9
6	-14.7

References

Arndt, Sandra et al. (Aug. 2013). "Quantifying the degradation of organic matter in marine sediments: A review and synthesis". en. In: *Earth-Science Reviews* 123, pp. 53–86. DOI: 10.1016/j.earscirev.2013.02.008.

Arnosti, C. (Dec. 2004). "Speed bumps and barricades in the carbon cycle: substrate structural effects on carbon cycling". en. In: *Marine Chemistry* 92.1-4, pp. 263–273. DOI: 10.1016/j.marchem.2004.06.030.

Beulig, F. et al. (Jan. 2018). "Control on rate and pathway of anaerobic organic carbon degradation in the seabed". en. In: *Proc. Natl. Acad. Sci. U.S.A.* 115.2, pp. 367–372. DOI: 10.1073/pnas.1715789115.

Canfield, Don E., Simon W. Poulton, and Guy M. Narbonne (Jan. 2007). "Late-Neoproterozoic Deep-Ocean Oxygenation and the Rise of Animal Life". en. In: *Science* 315.5808, pp. 92–95. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1135013. URL: https://www.science.org/doi/10.1126/science.1135013 (visited on 04/19/2024).

Close, H. G., R. Bovee, and A. Pearson (May 2011). "Inverse carbon isotope patterns of lipids and kerogen record heterogeneous primary biomass". en. In: *Geobiology* 9.3, pp. 250–265. ISSN: 1472-4677, 1472-4669. DOI: 10.1111/j.1472-4669.2011.00273.x. URL: https://onlinelibrary.wiley.com/doi/10.1111/j.1472-4669.2011.00273.x (visited on 04/18/2024).

Conrad, R (Mar. 1999). "Contribution of hydrogen to methane production and control of hydrogen concentrations in methanogenic soils and sediments". en. In: *FEMS Microbiology Ecology* 28.3, pp. 193–202. DOI: 10.1111/j.1574-6941.1999.tb00575.x.

Conrad, Ralf and Peter Claus (Aug. 2023). "Fractionation of stable carbon isotopes during microbial propionate consumption in anoxic rice paddy soils". en. In: *Biogeosciences* 20.17, pp. 3625–3635. DOI: 10.5194/bg-20-3625-2023.

Conrad, Ralf, Pengfei Liu, and Peter Claus (Dec. 2021). "Fractionation of stable carbon isotopes during acetate consumption by methanogenic and sulfidogenic microbial communities in rice paddy soils and lake sediments". en. In: *Biogeosciences* 18.24, pp. 6533–6546. DOI: 10.5194/bg-18-6533-2021.

DeNiro, Michael J. and Samuel Epstein (July 1977). "Mechanism of Carbon Isotope Fractionation Associated with Lipid Synthesis". en. In: *Science* 197.4300, pp. 261–263. DOI: 10.1126/science.327543.

DeVries, Tim (Oct. 2022). "The Ocean Carbon Cycle". en. In: *Annu. Rev. Environ. Resour.* 47.1, pp. 317–341. DOI: 10.1146/annurev-environ-120920-111307.

Drake, Travis W. et al. (Nov. 2015). "Ancient low–molecular-weight organic acids in permafrost fuel rapid carbon dioxide production upon thaw". en. In: *Proc. Natl. Acad. Sci. U.S.A.* 112.45, pp. 13946–13951. DOI: 10.1073/pnas.1511705112.

Franks, Stephen G et al. (Apr. 2001). "Carbon isotopic composition of organic acids in oil field waters, San Joaquin Basin, California, USA". en. In: *Geochimica et Cosmochimica Acta* 65.8, pp. 1301–1310. DOI: 10.1016/S0016-7037(00)00606-2.

Gilbert, Alexis et al. (Nov. 2009). "Accurate Quantitative Isotopic [13] C NMR Spectroscopy for the Determination of the Intramolecular Distribution of [13] C in Glucose at Natural Abundance". en. In: *Analytical Chemistry* 81.21, pp. 8978–8985. ISSN: 0003-2700, 1520-6882. DOI: 10.1021/ac901441g. URL: https://pubs.acs.org/doi/10.1021/ac901441g (visited on 05/27/2024).

Gonzalez-Nayeck, Ana C. et al. (May 2023). "Isotopic Signatures of Carbon Transfer in a Proterozoic Analogue Microbial Mat". en. In: *Applied and Environmental Microbiology* 89.5. Ed. by John R. Spear, e01870–22. ISSN: 0099-2240, 1098-5336. DOI: 10.1128/aem.01870-22. URL: https://journals.asm.org/doi/10.1128/aem.01870-22 (visited on 04/18/2024).

Gonzalez-Nayeck, Ana C. et al. (Sept. 2022). "Absence of canonical trophic levels in a microbial mat". en. In: *Geobiology* 20.5, pp. 726–740. ISSN: 1472-4677, 1472-4669. DOI: 10.1111/gbi.12511. URL: https://onlinelibrary.wiley.com/doi/10.1111/gbi.12511 (visited on 04/18/2024).

Hackmann, Timothy J. and Bo Zhang (Sept. 2023). "The phenotype and genotype of fermentative prokaryotes". en. In: *Sci. Adv.* 9.39, eadg8687. DOI: 10.1126/sciadv.adg8687.

Hayes, John M (2001). "Fractionation of the Isotopes of Carbon and Hydrogen in Biosynthetic Processes". en. In.

Heuer, Verena et al. (Nov. 2006). "Online δ ^{13} C analysis of volatile fatty acids in sediment/porewater systems by liquid chromatography-isotope ratio mass spectrometry: Online δ ^{13} C analysis of VFAs". en. In: *Limnol. Oceanogr. Methods* 4.10, pp. 346–357. DOI: 10.4319/lom.2006.4.346.

Heuer, Verena B., Fumio Inagaki, et al. (Dec. 2020). "Temperature limits to deep subseafloor life in the Nankai Trough subduction zone". en. In: *Science* 370.6521, pp. 1230–1234. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.abd7934. URL: https://www.science.org/doi/10.1126/science.abd7934 (visited on 04/19/2024).

Heuer, Verena B., Martin Krüger, et al. (Jan. 2010). "Experimental studies on the stable carbon isotope biogeochemistry of acetate in lake sediments". en. In: *Organic Geochemistry* 41.1, pp. 22–30. DOI: 10.1016/j.orggeochem.2009.07.004.

Heuer, Verena B., John W. Pohlman, et al. (June 2009). "The stable carbon isotope biogeochemistry of acetate and other dissolved carbon species in deep subseafloor sediments at the northern Cascadia Margin". en. In: *Geochimica et Cosmochimica Acta* 73.11, pp. 3323–3336. DOI: 10.1016/j.gca.2009.03.001.

Jørgensen, Bo Barker, Alyssa J. Findlay, and André Pellerin (Apr. 2019). "The Biogeochemical Sulfur Cycle of Marine Sediments". en. In: *Front. Microbiol.* 10, p. 849. DOI: 10.3389/fmicb.2019.00849.

Julien, Maxime et al. (Dec. 2022). "Re-evaluation of the 13C isotope fractionation associated with lipids biosynthesis by position-specific isotope analysis of plant fatty acids". en. In: *Organic Geochemistry* 174, p. 104516. ISSN: 01466380. DOI: 10.1016/j.orggeochem.2022.104516. URL: https://linkinghub.elsevier.com/retrieve/pii/S0146638022001504 (visited on 05/23/2024).

Klatt, Judith M et al. (Dec. 2020). "Versatile cyanobacteria control the timing and extent of sulfide production in a Proterozoic analog microbial mat". en. In: *The ISME Journal* 14.12, pp. 3024–3037. ISSN: 1751-7362, 1751-7370. DOI: 10.1038/s41396-020-0734-z. URL: https://academic.oup.com/ismej/article/14/12/3024-3037/7474743 (visited on 04/18/2024).

Lever, Mark A. et al. (Feb. 2010). "Acetogenesis in Deep Subseafloor Sediments of The Juan de Fuca Ridge Flank: A Synthesis of Geochemical, Thermodynamic, and Gene-based Evidence". en. In: *Geomicrobiology Journal* 27.2, pp. 183–211. DOI: 10.1080/01490450903456681.

Logan, Graham A. et al. (July 1995). "Terminal Proterozoic reorganization of biogeochemical cycles". en. In: *Nature* 376.6535, pp. 53–56. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/376053a0. URL: https://www.nature.com/articles/376053a0 (visited on 04/18/2024).

Lovley, D R (1991). "Dissimilatory Fe(III) and Mn(IV) reduction". en. In: *MICROBIOL. REV.* 55.

Lovley, Derek R. and Elizabeth J. P. Phillips (Apr. 1986). "Organic Matter Mineralization with Reduction of Ferric Iron in Anaerobic Sediments". en. In: *Appl Environ Microbiol* 51.4, pp. 683–689. DOI: 10.1128/aem.51.4.683-689.1986.

Monson, K. D. and J. M. Hayes (1980). "Biosynthetic control of the natural abundance of carbon 13 at specific positions within fatty acids in Escherichia coli. Evidence regarding the coupling of fatty acid and phospholipid synthesis." In: *Journal of Biological Chemistry* 255.23, pp. 11435–11441. DOI: https://doi.org/10.1016/S0021-9258(19)70310-X.

Moran, James J. et al. (May 2023). "Daylight-driven carbon exchange through a vertically structured microbial community". en. In: *Frontiers in Microbiology* 14, p. 1139213. ISSN: 1664-302X. DOI: 10.3389/fmicb.2023.1139213. URL: https://www.frontiersin.org/articles/10.3389/fmicb.2023.1139213/full (visited on 04/18/2024).

Mueller, Elliott P., Fenfang Wu, and Alex L. Sessions (Nov. 2022). "Quantifying Isotopologue Reaction Networks (QIRN): A modelling tool for predicting stable isotope fractionations in complex networks". en. In: *Chemical Geology* 610, p. 121098. DOI: 10.1016/j.chemgeo.2022.121098.

Neidhardt, Frederic Carl, ed. (1987). *Escherichia coli and salmonella typhimurium: cellular and molecular biology. 2*. en. Vol. 2. Washington, D.C: American Society for Microbiology.

Penning, Holger and Ralf Conrad (May 2006). "Carbon isotope effects associated with mixed-acid fermentation of saccharides by Clostridium papyrosolvens". en. In: *Geochimica et Cosmochimica Acta* 70.9, pp. 2283–2297. DOI: 10.1016/j.gca.2006.01.017.

Reay, Dave S. et al. (Oct. 2018). "Methane and Global Environmental Change". en. In: *Annu. Rev. Environ. Resour.* 43.1, pp. 165–192. DOI: `10.1146/annurev-environ-102017-030154`.

Rothman, Daniel H. (Jan. 2024). "Slow closure of Earth's carbon cycle". en. In: *Proc. Natl. Acad. Sci. U.S.A.* 121.4, e2310998121. DOI: `10.1073/pnas.2310998121`.

Stal, Lucas J and Roy Moezelaar (1997). "Fermentation in cyanobacteria". en. In: *FEMS Microbiology Reviews*.

Tang, T. et al. (Mar. 2017). "Geochemically distinct carbon isotope distributions in *Allochromatium vinosum* DSM 180 [T] grown photoautotrophically and photoheterotrophically". en. In: *Geobiology* 15.2, pp. 324–339. DOI: `10.1111/gbi.12221`.

Van Bodegom, Peter (May 2007). "Microbial Maintenance: A Critical Review on Its Quantification". en. In: *Microb Ecol* 53.4, pp. 513–523. DOI: `10.1007/s00248-006-9049-5`.

White, D., J.T. Drummond, and C. Fuqua (2012). *The Physiology and Biochemistry of Prokaryotes*. Oxford University Press.

Zhang, Xiaowen et al. (Aug. 2023). "Late acquisition of the rTCA carbon fixation pathway by Chlorobi". en. In: *Nature Ecology & Evolution* 7.9, pp. 1398–1407. ISSN: 2397-334X. DOI: `10.1038/s41559-023-02147-0`. URL: `https://www.nature.com/articles/s41559-023-02147-0` (visited on 04/17/2024).

Chapter 5

ISOTOPIC EVIDENCE OF ACETATE CYCLING IN THE PRECAMBRIAN CONTINENTAL CRUST.

Elliott P. Mueller[1*], Juliann Panehal[1], Min Song[2] Christian T. Hansen[3], Oliver Warr [2,4], Jason Boettger[5], Verena B. Heuer[3], Wolfgang Bach[3], Kai-Uwe Hinrichs[3], John M. Eiler[1], Victoria Orphan[1], Barbara Sherwood Lollar[2,6], Alex L. Sessions[1]

[1]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA

[2]Department of Earth Sciences, University of Toronto, Toronto, ON, Canada

[3]MARUM Centre for Marine Environmental Sciences, University of Bremen, Bremen, Germany

[4]Department of Earth Sciences, University of Ottawa, Ottawa, ON K1N 6N5, Canada

[5]Department of Earth, Environmental, and Resource Sciences, University of Texas at El Paso, El Paso, Texas

[6]Institut de Physique du Globe de Paris (IPGP), Université Paris Cité, 1 rue Jussieu, 75005, Paris, France
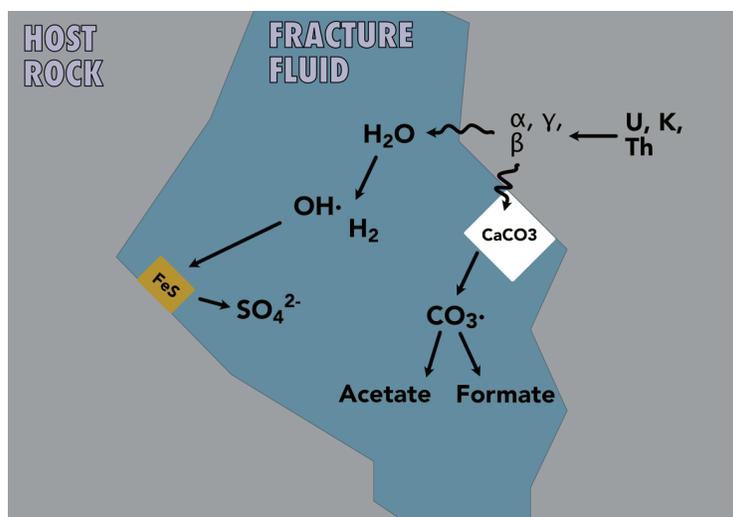
Graphical Abstract of Chapter 5

5.1 Abstract

The deep continental crust represents a vast potential habitat for microbial life where its activity remains poorly constrained. A common characteristic of these ecosystems is the presence of organic acids like acetate, but the role of these molecules in the subsurface carbon cycle - including the mechanism and rate of their turnover - is still unclear. Here, we developed an isotope-exchange 'clock' based on the temperature-dependent abiotic equilibration of H-isotopes between acetate's methyl-group and water, which can be used to define the maximum in situ residence time for acetate. We applied this technique to the fracture fluids in Birchtree and Kidd Creek mines within the Canadian Precambrian crust. At both sites, we found isotopic disequilibrium between acetate and water, indicating acetate residence times <1 million years and a rate of turnover that could theoretically support microbial life. However, radiolytic water-rock reactions could also contribute to acetate production and degradation, a process that would have global relevance for the deep biosphere. More broadly, our study demonstrates that isotope-exchange clocks can constrain in situ residence times of biomolecules with possible applications to other environments.

5.2 Introduction

Fluid-bearing fractures within crystalline rocks of the Precambrian continental crust have been identified globally at sites from the Canadian Shield to the South African Craton and may store as much as one-third of the Earth's groundwater (Ferguson et al., 2021). These fluids are generally characterized by anoxia, high salinities (up to 325 g/L), low cell densities ($< 10^3$ to 10^5 cells/L) and variable hydrogeologic

recharge rates (Bomberg et al., 2021; Holland et al., 2013; G. S. Lollar et al., 2019). At the Kidd Creek Cu-Zn-Ag Mine (Timmins, Ontario), noble gas-derived mean residence times of fracture fluids can exceed 10^9 years (Holland et al., 2013). Long fluid residence times allow the products of water-rock reactions to accumulate to a greater extent than elsewhere. Despite the accumulation of these potential substrates, cell density in the fluids is low, making the Kidd Creek Deep Fluid and Deep Life Observatory a prime window into abiogenic synthesis (G. S. Lollar et al., 2019). Most notably, radiolysis produces abundant H_2 while simultaneously generating oxidants like sulfate (Li-Hung Lin et al., 2005; Li-Hung Lin et al., 2005; B. S. Lollar et al., 2014; Warr, Song, and Barbara Sherwood Lollar, 2023). At sufficiently high concentrations, H_2 can reduce inorganic carbon to generate methane and higher hydrocarbons through abiotic Sabatier and polymerization reactions (B. Sherwood Lollar, Frape, et al., 1993; Etiope and Barbara Sherwood Lollar, 2013; Taguchi et al., 2022). It was recently suggested, based on laboratory experiments, that radiolysis in Kidd Creek may also generate simple organic acids such as acetate, formate and oxalate from water and dissolved inorganic carbon (Costagliola et al., 2017; Vandenborre et al., 2021; B. Sherwood Lollar, V.B. Heuer, et al., 2021). Indeed, the dissolved organic carbon pool in Kidd Creek's fracture waters is over 2 mM and up to 68% of this pool is composed solely of acetate and formate (B. Sherwood Lollar, V.B. Heuer, et al., 2021). Through observations of Kidd Creek and other subsurface continental sites, it has become clear that abiotic water-rock reactions including radiolysis can provide a chemical framework – organic carbon, oxidants and reductants – that could support microbial communities (Templeton and Caro, 2023).

The synthesis mechanism of these chemical species has been studied for over thirty years at Kidd Creek, yet estimates of their turnover times are to date limited. Methane and sulfur cycling have been examined through isotopic analyses, but these measurements provide binary statements about production and consumption rather than quantitative rates (Li et al., 2016; Warr, Giunta, et al., 2021). Substrate turnover times are instead estimated via bottom-up models of radiolytic yields that come with large uncertainties (Li-Hung Lin et al., 2005; Li-Hung Lin et al., 2005; B. S. Lollar et al., 2014; Warr, Song, and Barbara Sherwood Lollar, 2023). Direct measurements of carbon turnover are needed for accurate evaluation of the net productivity and thus habitability of hydrogeologically isolated systems like Kidd Creek. Moreover, environmental measurements of abiogenesis rates could elucidate the quantitative importance of these reactions in other deep biosphere locations both

on Earth and other planets or moons.

Here, we constrain the turnover time of acetate in two deep subsurface fracture fluid systems by developing and applying an isotope-exchange clock for dissolved acetate. First, we experimentally constrained the rate of uncatalyzed (abiotic) H-isotope exchange between water and acetate methyl-H, which is presumed to occur through a tautomerization reaction (Gao, 1996; Mardyukov, Eckhardt, and Schreiner, 2020). We found that the rate of this exchange reaction follows a first-order Arrhenius relationship with temperature. Since acetate is synthesized out of H-isotopic equilibrium with surrounding fluids and exchange drives it towards equilibrium at a known rate, the apparent ^{2}H-fractionation between acetate and water can serve as a clock: If acetate turnover is slower than abiotic isotopic exchange, acetate's methyl-site δ^2H composition will be defined by the water δ^2H and the equilibrium isotope effect (EIE) between them. Alternatively, if turnover is comparatively high, it will have a disequilibrated signature from the water. Although we do not (yet) know the magnitude of starting disequilibrium upon acetate synthesis, preventing a fully quantitative estimate of residence time, the mere presence of isotopic disequilibrium between acetate and water must indicate a residence time that is shorter than the equilibration time. We applied this approach to fracture fluids at Kidd Creek Mine and – for comparison – at Birchtree Mine, a more microbially active site in the Canadian Shield (B. Sherwood Lollar, V.B. Heuer, et al., 2021). At both sites, our data reveal acetate-water ^{2}H disequilibria and suggest relatively short acetate residence times (<1 Myr) in the Precambrian continental subsurface. These results also provide tentative constraints on the importance of radiolytic acetate production as an abiotic reaction in the deep biosphere.

5.3 Methods

Organic acid extraction. Organic acids were extracted following the procedure developed by Mueller et al. (2022) with minor changes to account for the high concentrations of chloride in the fracture waters. Briefly, samples of fracture fluid were titrated to pH >6 with NaOH if necessary. Samples were run through a Dionex Ag/H cartridge at 0.5 mL/min to remove chloride after the cartridge had been washed with 300 mL purified water (Milli-Q) at 2 mL/min. The first 0.5 mL of eluent from the cartridge was discarded as it represented the dead volume. The remaining sample was collected until almost all the resin was used, carefully avoiding over-filling the cartridge, which would cause chloride to leak through. The cartridge eluent was injected onto a Dionex high performance ion chromatography instrument with an

AG-11HC column and a KOH gradient from 1 to 20 mM. The organic acid fraction of the chromatogram was collected into vials using manual fraction collection. This step was repeated for samples with lower acetate concentration and collected into the same vial. The collected acids were titrated to pH >6 with degassed, anoxic NaOH and then dried down under nitrogen. Samples were redissolved in LC-MS grade methanol.

Stable isotope analysis. The majority of samples were analyzed on a heated electrospray ionization (HESI) Orbitrap QExactive HF (Thermo Fisher, Bremen, Germany) following the protocol of Mueller et al. (2022). This technique quantifies the molecular-average $\delta^{13}C$ (VPDB) and methyl-specific δ^2H of acetate by comparison to an external standard of sodium acetate ($\delta^{13}C$ = -19.2‰, d^2H = -127‰). Certain samples were measured on an electrospray ionization (ESI) Orbitrap Exploris 240, but the mass spectrometry parameters were identical and the same standard was used for all measurements. Multiple sample introduction methods into the Orbitrap were used throughout the course of this study.

Direct Infusion: A 500 uL syringe (Hamilton) was filled with sample or standard solution (in LC-MS grade methanol) and attached to a syringe pump (Chemyx). Solution was infused into the mass spectrometer at 5 uL/min. After a 7-minute acquisition, the syringe and its tubing were washed with 2 mL of LC-MS grade methanol and the next sample or standard was loaded into the syringe pump. This was repeated to achieve bracketed, sample-standard comparisons (AAAABBB-BAAAA, A = standard replicates, B = sample replicates). This method was used when memory effects between sample and standard due to large differences in d^2H or $d^{13}C$ were a concern. This was especially important for ^2H-enriched acetate samples from exchange experiments.

Dual Inlet: Two 500 uL syringes (Hamilton) were filled, one with sample and the other with standard solution (in LC-MS grade methanol) and attached to a syringe pump (Chemyx). Solution was infused into the mass spectrometer at 5 uL/min. Using a Rheodyne 6-port valve, sample and standard were alternated while achieving continuous flow of both (after Hilkert et al., 2021). Each acquisition block was 12 minutes with 4-5 minute switch times between blocks cut out of the data acquisition to avoid carryover effects. This was repeated to achieve bracketed, sample-standard comparisons (ABABABA, A = standard replicates, B = sample replicates). This method was used for the majority of Kidd Creek and Birchtree samples. Acetate standard was diluted to match sample ion current.

In-flow Injection: Samples were infused into the mass spectrometer using a Vanquish Horizons HPLC Split Sampler Autosampler and a Vanquish Horizons Pump set to 5 uL/min with degassed LC-MS grade methanol as an eluent. An injection volume of 50 uL was used to insert this sample into the flow of methanol which carried it to the Orbitrap for 14 minutes. At that time the flow rate was increased to 30 uL/min to clear residual sample from the transfer lines. At 18.5 minutes, the flow rate was dropped again to 5 ul/min and after 90 seconds, the next injection began. Data acquisition included all 20 minutes of the run but only integrated between 2 and 12 minutes to calculate isotope ratios. This was repeated to achieve bracketed, sample-standard comparisons (ABABABA, A = standard replicates, B = sample replicates). Acetate standard was diluted to match sample ion current.

In all of the above methods, the following ESI parameters were used as default. Minor adjustments were made daily to tune the instrument for spray stability. Polarity = negative, spray voltage = 3.0kV, spray current <0.2 uA, Auxiliary gas = 1 (arbitrary units), sweep gas = 1 (arbitrary units), sheath gas = 10 (arbitrary units), auxiliary gas temperature = 100°C, RF lens = 60%, capillary temperature = 320°C. The following Orbitrap parameters were used for all analyses. Automated gain control = 1e6, resolution = 60,000, microscans = 1, quadrupole range = 57 - 62 m/z, lock mass = off. Raw data off the Orbitrap was extracted using the software IsoX (Thermo Fisher, Bremen, Germany) and converted to isotope ratios using a Python script. This script uses the Makarov equation outlined in Mueller, Wu, and Sessions, 2022 to convert from ion intensities to ion counts. It then culls scans that are >99th percentile or <1st percentile in total ion current to avoid integrating scans with ion source aberrations.

Exchange Reactions

High temperature acetate-water exchange experiments were conducted using a customized Dickson-type flexible reaction cell setup (Parr Instruments) with no vapor phase present. Each flexible gold bag was filled with 90 mL of 1 mM sodium acetate in Milli-Q water (pH 6-7) that was sparged with nitrogen and pressurized to 30 MPa. Two experiments were performed at 150°C in 5% ^2H_2O. One was run for a week, sampling every 24 hours, while the other was run for a month, sampling every 3-5 days. Another month-long experiment with 5% ^2H_2O was performed at 100°C, sampling every 3-5 days. Acetate-water exchange experiments were also performed at 60°C in 60 mL serum vials. Each vial was filled with 50 mL of 1 mM sodium acetate in 5% $2H_2O$ (pH 7) that had been sparged with nitrogen and sealed with a

butyl rubber stopper and crimped with an aluminum cap. At each timepoint, 1 mL of sample was collected via needle and syringe. The sample was immediately frozen and stored at -20°C and the solution was sparged with nitrogen again to remove any air introduced during sampling. These experiments were done in triplicate. All exchange experiments were performed at pH 6-7 to match environmental conditions.

Additional high temperature flexible gold bag experiments were performed to determine the equilibrium isotope effect at 200°C (30 MPa). Each reaction cell was filled with 90 mL of 1 mM sodium acetate (pH 6-7) in either -50‰ or +110‰ $\delta^2 H$ water. Each condition was measured in duplicate, resulting in four total experiments. Samples were taken every hour for the first six hours to measure the extent of isotopic exchange with time and then every ~6-12 hours for the next 66 hours. At each time point, 1.5 mL of the sample was collected and discarded to remove the dead volume from the sampling apparatus and then an additional 1.5 mL of sample was taken for acetate $\delta^{13} C$ and $\delta^2 H$ analyses. Collected aliquots were immediately frozen and stored at -20°C until they were analyzed.

The kinetic rate constants for H-isotope exchange were calculated using the formulation from Sessions et al., 2004:

$$\frac{F_e - F_t}{F_e - F_i} = e^{-kt} \tag{5.1}$$

where F_t is the ^{2}H fractional abundance (i.e. mole fraction) at a given timepoint, F_i is the initial fractional abundance and F_e is the fractional abundance at equilibrium. The latter was calculated using the fractional abundance of the water and the equilibrium isotope effect from DFT models at the corresponding temperature. In experiments where the isotope composition approaches or reaches equilibrium, data points close to the equilibrium value were discarded from the calculation of rate constant due to the large propagated errors when the natural logarithm of the value $F_e - F_t$ was close to zero.

Isotope Fractionation Calculations.

The apparent hydrogen isotope fractionation between acetate and water ($\epsilon_{\text{acetate/water}}$) was calculated as:

$$^2\alpha_{acetate/water} = \frac{\delta^2 H_{Acetate} + 1000}{\delta^2 H_{Water} + 1000} \tag{5.2}$$

$$^2\epsilon_{acetate/water} = (^2\alpha_{acetate/water} - 1) \times 1000 \tag{5.3}$$

Thermodynamics and cell density calculations.

The free energy (DG) available to microbial metabolisms was calculated by adjusting the standard free energy (DG°) for the activity of the reactants and products found in Kidd Creek fracture fluids following the equation:

$$\Delta G = \Delta G° + RTlnQ \tag{5.4}$$

where R is the ideal gas constant (kJ/mol/K) and T is temperature (K), set to 298K at 500 bar pressure. Q is the reaction quotient defined as:

$$Q = \prod a_i^{v_i} \tag{5.5}$$

where a is the activity of a substrate defined as the product of its concentration (molar) and gamma value and v is the stoichiometric coefficient which is negative for reactants. Gamma values for sulfate, methane and bicarbonate were found on the Geochemists Workbench with the thermo-hmw.dat database, which uses a Pitzer equation based Harvie-Møller-Weare activity model owing to the high ionic strength of the fracture fluid (4.9 molal). Acetate is not part of this database, so it was calculated with extended Debye Hueckel equation using the thermo.dat database. The concentrations used in these calculations were taken from data in Lin et al. (2016). Sulfate, bicarbonate, acetate, and methane concentrations were set to 620 mM, 57 mM, 1.3 mM, and 2.1 mM, respectively. Methane concentration was calculated from fluid flow rate, gas exsolution rate from the fluid, and the concentration of methane in the gas (from Lin et al. 2016). It was assumed that all methane was dissolved fully in solution due to the high (500 bar) *in situ* pressure of the fracture fluids (after Sherwood-Lollar et al. 1993). Sulfide was below detection limits (< 2 mM). Its concentration was set to 10 nM but increasing its concentration to the detection limit did not change the implications of the cell densities (>10 cells/mL at all maintenance energies simulated).

Cell density (cells/L) is calculated by combining the acetate turnover rate (M/s), the free energy of the reaction (J/mol), and the maintenance energy of a cell (J/s/cell).

$$\rho = \frac{\tau_{AC} \times \Delta G}{ME} \tag{5.6}$$

where t_{AC} is the turnover time and r is the cell density.

Density functional theory calculations of EIE.

Temperature-dependent ^{2}H/^{1}H equilibrium fractionation between acetate and water was estimated using density functional theory. Liquid-phase acetate and water molecular models were optimized in the GAUSSIAN(TM) program, revision D.01[43] using basis set 6-311G(d,p)[43,44] and functional B3LYP under Tight optimization criteria (maximum/RMS atomic displacement 0.00006/0.00004 Bohr, maximum/RMS force 0.000015/0.00001 Hartrees/Bohr or Hartrees/Radian), with an Ultrafine integration grid mesh (Stephens et al., 1994; Lee, Yang, and Parr, 1988). The integral equation-formalism polarizable continuum model was used to represent the solvation environment. Following optimization, frequency calculations were carried out for the monoisotopic isotopologues and with a single D/H substitution to determine the effect of D/H substitution on vibrational frequencies. The Urey-Bigeleisen-Mayer equation was used to calculate the temperature-dependent reduced partition function ratio of each species under D/H substitution (Urey, 1947). Corrected ratios were computed using the temperature-dependent regression of Wang et al. (2009) to account for the effects of anharmonicity. The equilibrium fractionation factor was then computed as the ratio of the corrected ratios at the desired temperature.

Radiolytic yield calculations

To estimate the radiolytic yield (nM/J) of acetate production by alpha, gamma, and beta irradiation in Kidd Creek needed to support a given rate of acetate production, modified calculations from Warr et al. (2023) were used. The total acetate yield (Y_{AC}) in nM/s is defined as:

$$Y_{AC} = \frac{\sum E_{net,i} \times G_i \times \rho_{bulk}}{\phi} \tag{5.7}$$

where i represents either alpha, gamma or beta radiation and E_{net} is the dose rate (Gy/s) and G is the radiolytic yield (G). The bulk rock density (r_{bulk}) was set to 2.98 kg/dm^3. ϕ is the porosity, typically ~1% at crystalline rocks sites like Kidd Creek (Warr, Song, and Barbara Sherwood Lollar, 2023). Here, we assume that beta and gamma radiation does not produce acetate, since it has not been measured, such studies have not yet been done, so only α radiation is considered. Consequently, this represents a conservative estimate of radiolytic acetate production. Alpha radiolytic yields were taken from Vandenborre et al., 2021. In experiments with 200 mM dissolved carbonate in pure water, acetate accumulated to 8 mM within 1400 Gy

of absorbed radiation and plateaued at this concentration up to 5600 Gy, due to competing production and consumption reactions reaching a steady state. This results in a range of 1.3 to 6.0 nM/J for alpha radiation yields.

The dosage rate of alpha radiation is calculated as:

$$E_{net,\alpha} = \sum \frac{E_{\alpha,X} \times W \times S_\alpha}{1 + W + S_\alpha} \qquad (5.8)$$

Where E_α is the dosage of alpha radiation emitted (Gy/s) and X represents the specific elemental source of that radiation. S_α is the stopping power of rock to alpha radiation set at 1.5 after Warr et al. (2023). W is the water-rock ratio set to 0.37% calculated following Warr et al 2023, using water and rock density of 1.11 g/cm^3 and 2.98 g/cm^3, respectively, and a porosity value of 1% (Li-Hung Lin et al., 2005; Warr, Song, and Barbara Sherwood Lollar, 2023).

At 1% K , 1 ppm Th, and 1 ppm U, these elements emit 0, 1.93 x 10^{-12} and 6.9 x 10^{-12} Gy/s of alpha radiation, respectively (Warr, Song, and Barbara Sherwood Lollar, 2023). To estimate E_α for each of these elements in Kidd Creek, they were linearly increased based on the actual concentration in the deposit, which are 1.5 ppm, 6.7 ppm and 1.7% for U, Th and K, respectively (Warr, Song, and Barbara Sherwood Lollar, 2023). Therefore, the E_α for U, Th, and K is estimated at 0, 1.3 x 10^{-11} and 1.0 x 10^{-11} Gy/s in Kidd Creek.

5.4 Results

Experimental rates of hydrogen isotope exchange between acetate methyl hydrogen and water

Acetate was incubated at temperatures between 60°C and 200°C in the presence of 5% deuterated water in pressurized gold bags (see Methods). To derive the kinetic rate constant for hydrogen exchange between acetate's methyl group and ambient water, the ^{2}H/^{1}H ratio (δ^2H value) of acetate's methyl group was measured periodically throughout the incubations via ESI-Orbitrap mass spectrometry (See Methods) (Mueller, Sessions, et al., 2022). Under every condition tested, acetate δ^2H values increased with time reflecting exchange with the ^{2}H-enriched aqueous medium. At high temperatures (>150°C), the rate of acetate ^{2}H enrichment over time was initially linear then gradually flattened as it approached isotopic equilibrium with water (Supplementary Figure 5.6). At lower temperatures, the exchange kinetics were too slow to allow full equilibration of acetate and water within the runtime of the exper-

iments. The fitted half-times for exchange increased exponentially with decreasing temperature from 3 hours to 810 years, following an Arrhenius relationship (R^2 = 0.999, E_A = 138 kJ/mol, Figure 5.1). Replicate incubations, which were performed for all conditions except 100°C, resulted in similar reaction rates (overlapping data points in Figure 5.1, Supplementary Table 3). Exchange between acetate's methyl-site and water is presumed to occur through a reversible tautomerization between ethanoate and ethenol moieties (Supplementary Figure 5.8). Regardless of the exact mechanism, the excellent fit to an Arrhenius relationship between 60°C and 200°C suggests that the mechanism of exchange does not change within the tested temperature range. Extrapolating to the ambient temperature for samples collected at Kidd Creek and Birchtree (25°C), the predicted exchange half-time was 250,000 ± 70,000 years (2xRMSD).

Determining the equilibrium isotope effect

Equilibrium ^{2}H-isotope effects (EIEs) for acetate-water were calculated using density functional theory (DFT) across a range of temperatures (see Methods). These indicated a temperature-dependent change in the EIE from -108‰ at 250°C to -192‰ at 25°C (Figure 5.2B). Four high temperature incubations at 200°C were designed to experimentally test these calculations. Incubations were started with varying magnitudes and direction of isotopic disequilibrium, but in each case acetate δ^2H values changed with time until the experiments converged to similar EIEs. Water was present in excess and so did not change in δ^2H value. Equilibrium was reached in less than one day at 200°C and remained there for two days (Figure 5.2A). On average, the measured EIE (0.888 ± 0.012) was within analytical error of the DFT-calculated value (0.882). While the two experimental series did not perfectly converge in δ^2H values, they came within ~20‰ of each other. This offset is potentially due to analytical artifacts associated with measuring the high δ^2H value of acetate in the ^{2}H$_2$O spiked sample and is small in comparison to the scale of natural hydrogen isotope variations (blue, Figure 5.2A). Thus, at 200°C, the empirically determined EIE corroborates the DFT calculations.

Carbon and hydrogen isotope compositions of acetate from deep mines.

The $\delta^{13}C$ and δ^2H values of acetate extracted from Kidd Creek and Birchtree fracture fluids were measured via the ESI-Orbitrap method, revealing different isotopic compositions at the two sites (Mueller, Sessions, et al., 2022). Samples collected from three separate boreholes in Kidd Creek between 2008 and 2018 yielded $\delta^{13}C$ values of -10.0‰ to -6.6‰ (VPDB) and δ^2H values of -142‰ to
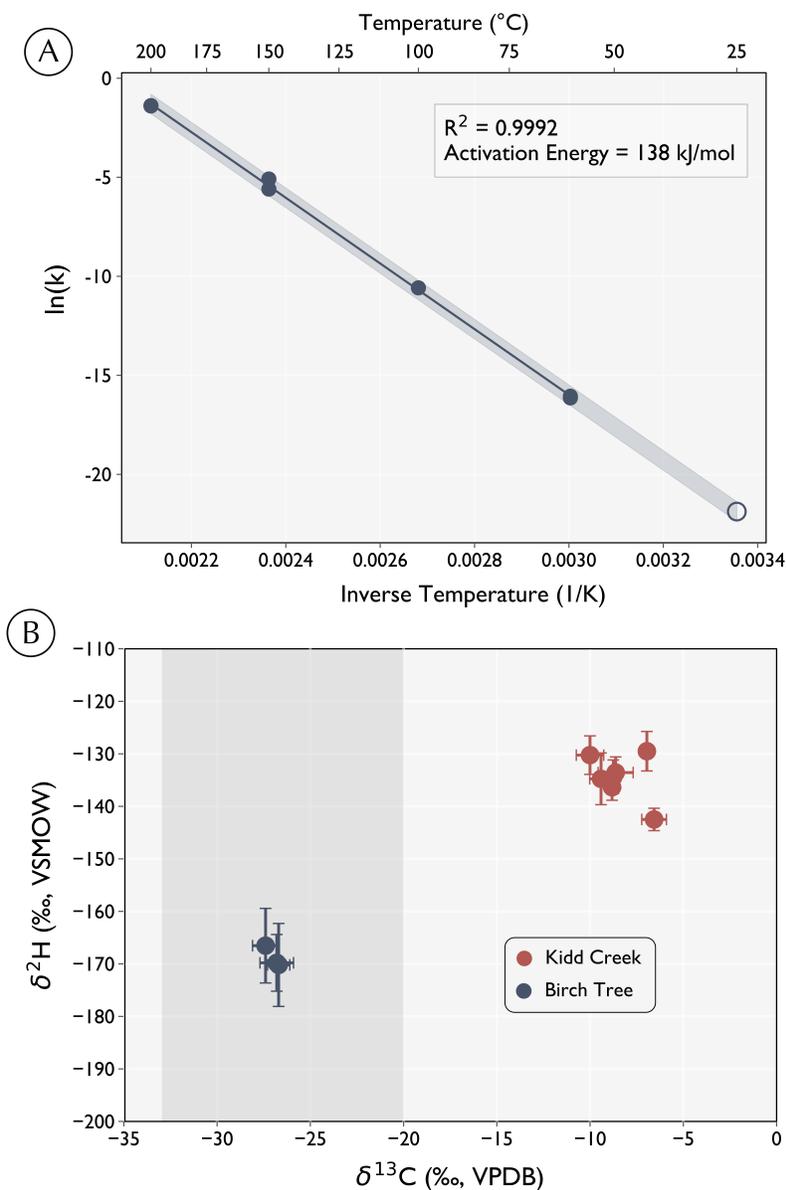
Figure 5.1: (A) Arrhenius plot of hydrogen isotope exchange rates with a linear regression through experiments at 60°C (n=3), 100°C (n=1), 150°C (n=2), and 200°C (n=2) (solid circles). Extrapolated reaction rates are projected to 25°C (open circle). Shaded region represents 2 RMSD. (B) Carbon and hydrogen isotope composition of acetate from Kidd Creek and Birchtree mines. Shaded regions represent $\delta^{13}C$ of total organic carbon from the metasedimentary rocks of the Kidd Creek formation (Blaser, Dreisbach, and Conrad, 2013).

-130‰ (VSMOW). In contrast, acetate extracted from three fracture fluid samples from Birchtree yielded $\delta^{13}C$ values of -26.7‰ to -27.4‰ and δ^2H values of -167‰ to -170‰ (Figure 5.2, Supplementary Table 2). All $\delta^{13}C$ values match the range of
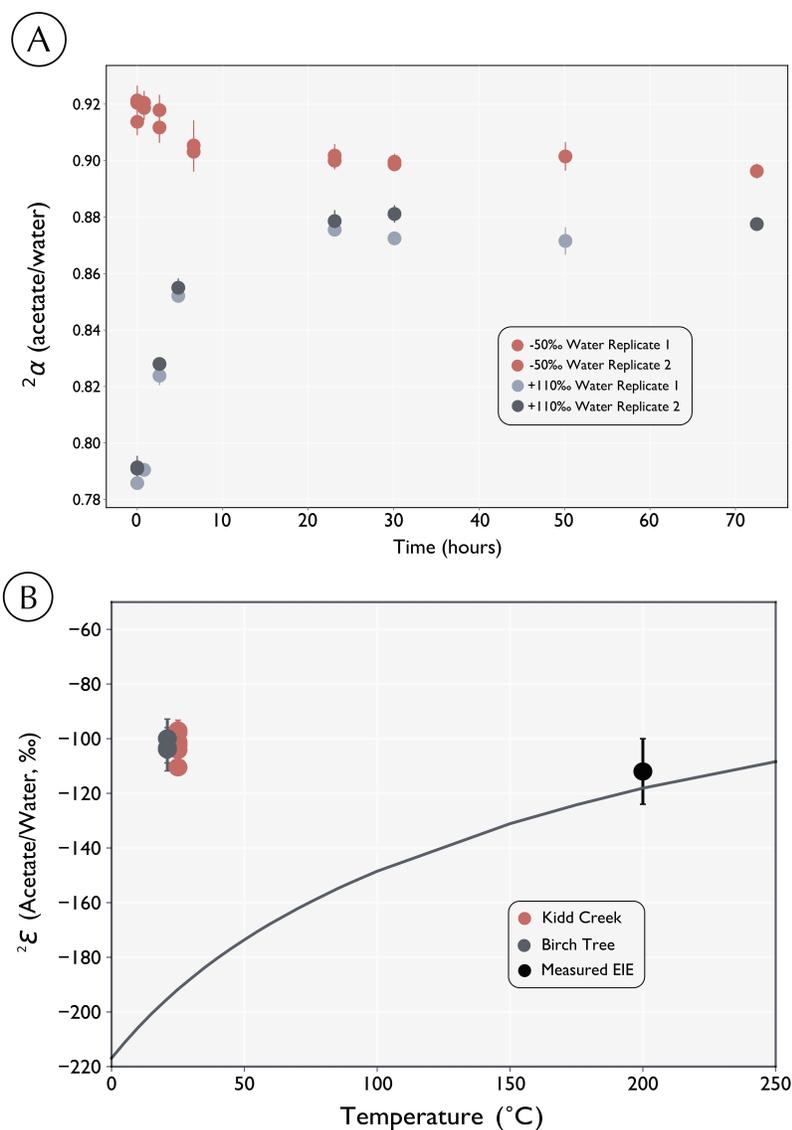
Figure 5.2: (A) Observed isotope effect between acetate and water throughout a three-day 200°C exchange experiment with water at either -50‰ or +110‰. (B) Hydrogen isotope fractionation between acetate and water ($^2\epsilon_{acetate/water}$) at both sites and empirically-derived equilibrium isotope effect from high-temperature exchange experiments. Solid line is the calculated isotope effect as a function of temperature.

values previously reported for these two sites (B. Sherwood Lollar, V.B. Heuer, et al., 2021). When compared to the previously-measured δ^2H values of water from Kidd Creek and Birchtree (-36‰ and -74‰, respectively), a similar apparent hydrogen isotope fractionation between acetate and water exists at both sites (Warr, Giunta, et al., 2021). This fractionation ranges from -115‰ to -90‰ (Figure 5.2B) and

differs from isotopic equilibrium at 25°C by nearly 100‰. These data demonstrate that acetate in Kidd Creek and Birchtree fracture fluids is far from the calculated H-isotopic equilibrium with water and must therefore have rates of production and consumption that are faster than the rate of abiotic exchange.

The identical apparent acetate-water hydrogen isotope effect ($^2\epsilon_{\text{acetate/water}}$) from the two sites is notable (Figure 5.2B). One possibility that we considered is whether complexation of acetate by the abundant (>1M) dissolved cations could significantly alter the EIE, i.e. a 'matrix effect' (G. S. Lollar et al., 2019). In this case, a shared $^2\epsilon_{\text{acetate/water}}$ value between the sites would be feasible if acetate at both sites was in equilibrium with water and the $^2\epsilon_{\text{acetate/water}}$ value matched the shifted EIE. To our knowledge, the speciation of acetate in salt mixtures and the resulting isotope effects are not constrained. However, we contend that complexation is unlikely to shift the EIE for methyl-group H by ~100‰ given that the hydrophobic methyl group would not participate in complexation with ions. Thus, all available data suggest that acetate and water in Kidd Creek and Birchtree are in substantial isotopic disequilibrium and the identical value of $^2\epsilon_{\text{acetate/water}}$ values observed at both sites (Figure 5.2B) likely reflects kinetic isotope effects that provide insight into acetate turnover mechanisms.

5.5   Discussion

The turnover times of organic molecules can provide important constraints on the productivity and habitability of isolated systems like the continental deep biosphere, but to date such timescales have been difficult to measure (Templeton and Caro, 2023). Water-rock reactions influencing the geochemistry of Kidd Creek and other sites often operate too slowly to replicate through experimentation. Similarly, microbial growth rates and metabolic fluxes typical of these settings are inaccessibly slow on laboratory timescales (Bomberg et al., 2021). While these processes can be identified through isotope geochemistry and genomic analyses, rates of abiogenesis and/or microbial metabolism remain elusive (Magnabosco et al., 2016). Our new H-isotope exchange clock helps to fill that gap by setting upper limits on residence times (i.e. lower limits on production and consumption rates) for acetate. Moreover, the general approach should be directly applicable to other organic molecules in the environment including many potentially important organic substrates and biomolecules.

Acetate is cycled in the continental deep subsurface.

In fracture fluids from both Kidd Creek and Birchtree, isotopic disequilibrium

between acetate and water implies active production and consumption of acetate by physical, chemical, and/or biological processes. These processes must generate and consume acetate faster than the abiotic exchange reaction can establish H-isotope equilibrium with water. Given that equilibration of hydrogen atoms occurs in less than four half-times, acetate residence times must be less than one million years, at least 1000-fold shorter than that of the host fracture fluids (>1 Gyr). Normalizing by the concentrations of acetate (Supplementary Table 2) and assuming present-day concentrations are at steady-state, these turnover times require acetate production and consumption rates of >1 nM/year and >0.1 nM/year in Kidd Creek and Birchtree, respectively. Since estimated physical fluid recharge rates are slower than acetate turnover times, our data suggest active production and consumption of acetate by microbial metabolisms and/or abiotic reactions (Holland et al., 2013).

Acetate consumption could support microbial communities.

Many anaerobic microorganisms use acetate as a carbon and electron source. The rates of acetate consumption implied by our residence time estimates provide an opportunity to quantify the amount of metabolic power potentially available to microbes consuming this substrate in the continental deep biosphere. Anaerobic respiration – represented here as sulfate reduction – and methanogenesis are common acetate consumption pathways in anoxic environments (Jørgensen, Findlay, and Pellerin, 2019; Blair and Carter, 1992). Considering the lower threshold of 1 nM/year for acetate consumption in Kidd Creek, acetate would supply $10^{-9.6}$ W/L or $10^{-9.8}$ W/L via sulfate reduction or methanogenesis, respectively (Figure 5.3). Assuming a range of cell-specific maintenance powers (the flux of energy required to maintain a cell), this rate could support between 10^2 to 10^6 cells/mL (Figure 5.3) (Tijhuis, Van Loosdrecht, and Heijnen, 1993; Tappe et al., 1999; Bradley et al., 2020). In saline fracture fluids of the continental subsurface, microbial cells must synthesize organic osmolytes to combat high osmotic pressures, increasing their basal power demands (Telling et al., 2018; Van Bodegom, 2007). Our results suggest that even with these higher power requirements, at least 10^3 cells/mL could theoretically survive solely on acetotrophic metabolic pathways in Kidd Creek (Figure 5.3). However, such calculations only reveal the viability of these prospective metabolic pathways and cannot be used as sole evidence of microbial acetotrophy. Further evidence is required to determine whether acetate is actively being consumed by biotic processes.

Constraining acetate sources and sinks in the subsurface
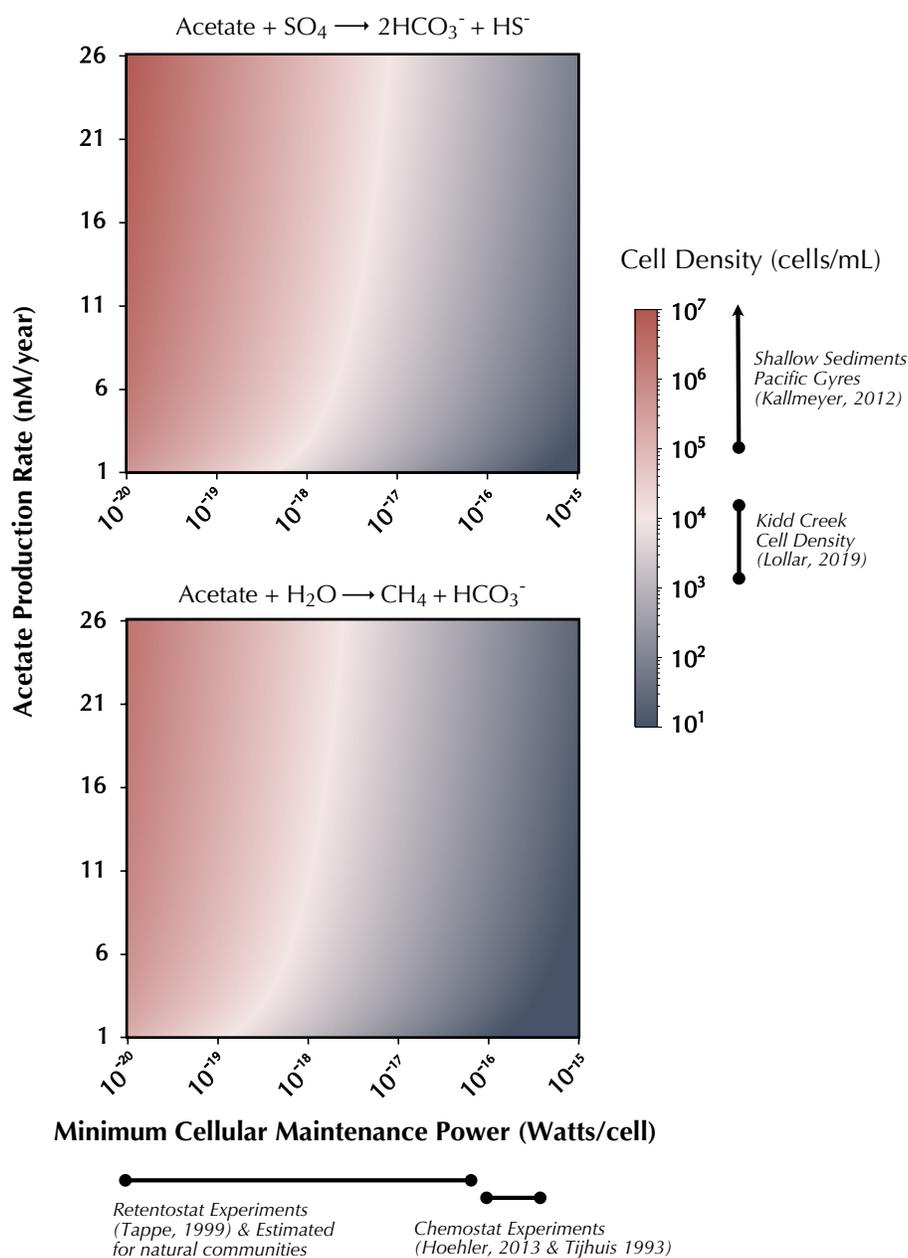
Figure 5.3: Theoretical cell densities for sulfate reducers (left) and acetoclastic methanogens (right) that could be supported in the fracture fluids, using a range of acetate production rates and minimum cellular maintenance power.

Microbial metabolisms:

The processes producing and degrading acetate can be constrained using its steady-state isotopic composition. In anoxic settings, acetate typically has a $\delta^{13}C$ value similar to that of the surrounding total organic carbon (TOC). This is commonly

attributed to minimal isotope effects associated with the production of acetate by microbial fermentation and consumption by anaerobic respiration (V. B. Heuer et al., 2009; Thomas, 2008; J T Gelwicks, J. B. Risatti, and J. M. Hayes, 1994). Acetate in Birchtree fracture fluids has $\delta^{13}C$ values that match this expectation, but it does not have the characteristic ^{13}C and ^{2}H depletion associated with chemolithoautotrophic acetogenesis (Mueller, Sessions, et al., 2022; Jeffrey T. Gelwicks, J. Risatti, and J. Hayes, 1989; Lever et al., 2010; Blaser, Dreisbach, and Conrad, 2013). This suggests that acetate turnover in Birchtree fluids is driven by heterotrophic microbial metabolisms.

In contrast, acetate in Kidd Creek is ^{13}C-enriched relative to TOC (Wellmer et al., 1999). If microbial activity is similarly responsible for acetate turnover in Kidd Creek fracture fluids, the reactions(s) consuming acetate must have larger (normal) carbon isotope effects than those in Birchtree. Acetoclastic methanogenesis exhibits such an isotope effect (25-30‰) (Warr, Giunta, et al., 2021). When fermentation of organic matter to acetate is coupled with methanogenic consumption, acetate can indeed be ^{13}C-enriched relative to TOC; however, this enrichment is not consistent across environments and the mechanisms behind it are still unclear (Blair and Carter, 1992; Verena Heuer et al., 2006; Thomas, 2008; J T Gelwicks, J. B. Risatti, and J. M. Hayes, 1994). Furthermore, the isotopic composition of methane and low ratio of methane-to-higher-alkanes in Kidd Creek fluids are not consistent with the significant rates of acetoclastic methanogenesis required to generate the observed ^{13}C enrichment in acetate (Warr, Giunta, et al., 2021; B. Sherwood Lollar, Frape, et al., 1993). As such, other mechanisms should be considered to explain acetate turnover in this system.

Radiolytic reactions:

The identical hydrogen isotope fractionations between acetate and water at Kidd Creek and Birchtree could indicate turnover mechanisms that are shared between the mines, such as radiolytic reactions. Radiolysis is well documented in the deep biosphere and has been shown to both produce and degrade acetate in laboratory experiments (Costagliola et al., 2017; Vandenborre et al., 2021). Radiolytic reactions occur when alpha, beta and gamma irradiation from natural decay of U, Th and K in the rock matrix triggers reactions with surrounding water, solutes, and minerals (Nisson et al., 2023; Blaser, Dreisbach, and Conrad, 2013). Since radiolysis drives substantial abiotic chemistry in subsurface fluids (i.e. H_2 production), it could produce acetate *in situ* as well (Li-Hung Lin et al., 2005; B. S. Lollar et al., 2014;

Costagliola et al., 2017; Vandenborre et al., 2021; B. Sherwood Lollar, V.B. Heuer, et al., 2021).

If radiolytic synthesis is the source of acetate in these fracture fluids, it operates at a rate that far exceeds those observed in laboratory studies. Maximum net yield during *in vitro* experiments is 6 nM acetate per joule of alpha radiation, corresponding to 0.007 nM/yr acetate generation rate in Kidd Creek fluids (see Methods), well below the minimum production rate estimated here (Vandenborre et al., 2021). These results should be interpreted with caution though. Radiolytic synthesis of organic acids is not a single production reaction but a network of reactions that both creates and degrades acetate (Vandenborre et al., 2021; Costagliola et al., 2017). The net yield measured *in vitro* represents a balance of production and degradation fluxes whereas gross yields could be much higher. Thus, if radiolysis is both producing and degrading acetate *in situ,* it could support fast turnover times without having high net generation rates. Kinetic isotope effects associated with this turnover could then explain the constant hydrogen isotope fractionation from water observed at both sites. However, while radiolysis is likely cycling acetate in the continental subsurface to some extent, we cannot presently determine whether it is *solely* responsible for acetate turnover based on our current understanding.

Future work should carefully examine radiolytic reactions under conditions that match the subsurface to assess their rates of acetate turnover and associated isotope effects. Given that the substrates for radiolysis – water and DIC – are ubiquitous, this process could provide a means to fuel acetotrophic metabolisms in environments well beyond the Precambrian continental subsurface, including global marine sediments, groundwaters, and the subsurface of other planets or moons.

Isotope-exchange clocks may have wide-ranging applications

Isotope-exchange clocks may also be relevant for other molecules and environments. For isolated systems characterized by slow turnover (i.e. subsurface environments of Earth, Mars or Europa), the acetate H-exchange reaction introduced here could be a useful constraint on acetate residence times or could simply confirm the presence of an active carbon cycle. However, for more biologically productive environments with fast turnover of organics (i.e. shallow marine sediments), this particular clock is insensitive (Beulig et al., 2018). Isotope exchange in organic molecules that experience more rapid equilibration of C-bound H would provide more useful information about substrate turnover in these systems. Molecules containing acidic alpha-H atoms, which can undergo tautomerization more easily than acetate (i.e.

longer chain organic acids and aldehydes), are potential targets (Ouellette and Rawn, 2015). Conversely, molecules with yet slower exchange (e.g. alkanes) could provide information about turnover in hotter environments (Hilkert et al., 2021). Our study provides the analytical and experimental basis for developing these techniques and directly constraining the turnover of small biomolecules *in situ* using their hydrogen isotope composition, one that could be applied to diverse environments.

5.6   Supplementary Materials

Acetate was purified from the high salinity matrix of fracture fluids prior to isotopic analysis. To verify that the purification scheme did not introduce isotopic fractionation, we spiked synthetic solutions of Kidd Creek fluid (Table S1) with acetate of known isotopic composition ($\delta^{13}C$ = -19.2‰, δ^2H = -127‰) and then reanalyzed them after extraction from the solution. Replicates of 1 mM and 2 mM acetate solutions yielded identical $\delta^{13}C$ and δ^2H values within analytical uncertainty, indicating that the extraction procedure is not fractionating. (Figure 5.4).

To account for changes in exchange reaction rates due to complexation of acetate with inorganic cations in solution, acetate was incubated in 1 mol/L $CaCl_2$ for 4 hours at 120°C. This experiment was done in triplicate. After the incubation, acetate was enriched by 191‰ ± 4‰. Uncertainty represents standard deviation of experimental replicates, which was equivalent to analytical error. This enrichment with time was converted to a reaction rate of exchange at 120°C, which falls on the established Arrhenius relationship between temperature and reaction rate based on exchange with pure water. These data suggest that complexation with calcium cations, the major cation in Kidd Creek, does not impact exchange kinetics.

Table 5.1: Composition of synthetic solution used for validation studies on acetate purification techniques.

Component	Concentration (g/L)
$CaCl_2$	144.3
NaCl	36.8
KCl	0.2
$MgCl_2$	8.4
NaBr	2.5
$NaNO_3$	0.02
$NaNOl_2$	0.01
KHl_2POl_4	0.03
$MgSOl_4$	0.07
Na-formate	0.03

The tautomerization reaction that putatively exchanges hydrogen atoms with water is shown below (Figure 5.7). Tautomerization reactions are well-documented iso-merizations between carbonyl/carboxyl and enol forms of organic ketones and acids (Mardyukov, Eckhardt, and Schreiner, 2020; Ouellette and Rawn, 2015). The good agreement with the Arrhenius equation between 60°C and 200°C suggests that the reaction mechanism of exchange does not change in this temperature range. Experiments at lower temperatures are not possible on laboratory timescales due to the long time for exchange (thousands of years). At the in situ temperature of Kidd Creek and Birchtree (~25°C), the extrapolated half-time of exchange is 250,000 years. If the relationship between temperature and rate were to deviate from the regression at lower temperatures, it would represent a shift in exchange mechanism. This new mechanism would have to have a faster rate at 25°C than the one measured at high temperatures, otherwise this mechanism would not be rate limiting. If this new mechanism was slower, the estimated rate of exchange by tautomerization would not be impacted, as it would still be rate limiting. This means that the exchange rate is a conservative, minimum estimate. If another mechanism was rate-limiting, our extrapolation would be an overestimate of the exchange half-time in Kidd Creek and Birchtree, and faster exchange rates only further the conclusions drawn in the following sections regarding acetate cycling rates.
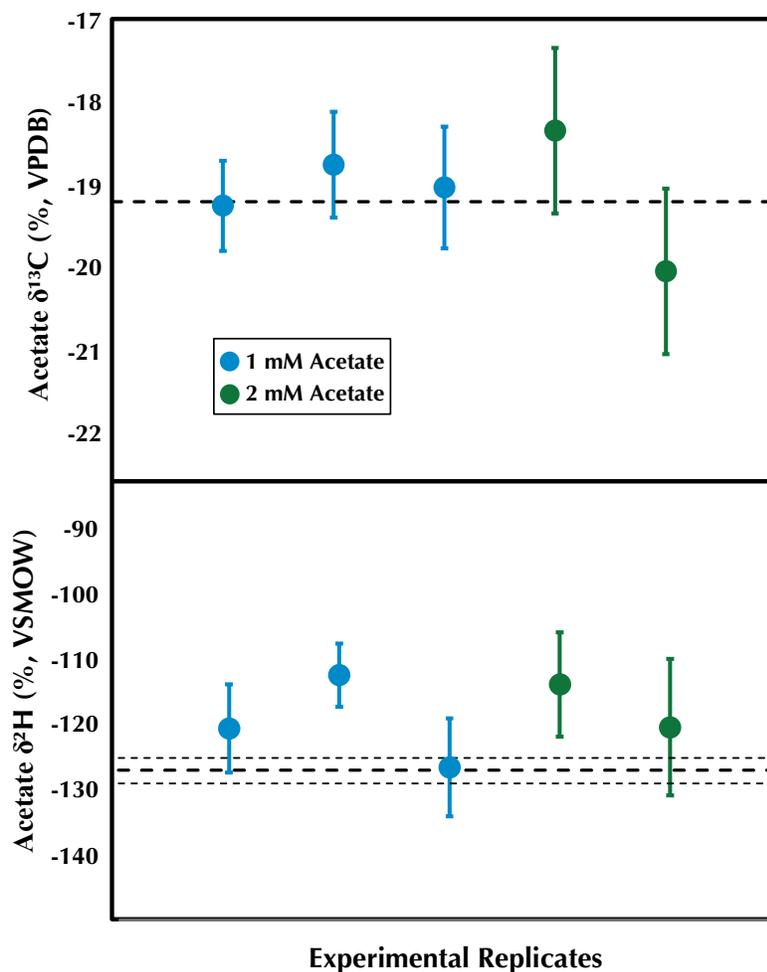
Figure 5.4: Extracted acetate standard from a synthetic fracture fluid mixture has the same $\delta^{13}C$ and $\delta^2 H$ values as the known composition. Error bars are one standard deviation on triplicate analytical replicates. Dotted lines represent the reported value of the standard along with uncertainties on $\delta^2 H$ values. Uncertainties are 0.1‰ on the reported $\delta^{13}C$ composition and are not visible on this plot.

Table 5.2: Composition of synthetic solution used for validation studies on acetate purification techniques.

Experiment Date	Temperature (°C)	Half-time (hr)	Half-time (yr)
February 2023	60	7.1×10^6	811
February 2023	60	6.5×10^6	742
June 2022	100	2.7×10^4	3.08
April 2022	150	112	0.012
June 2022	150	187	0.021
January 2023	200	3	0.00034
January 2023	200	3	0.00034
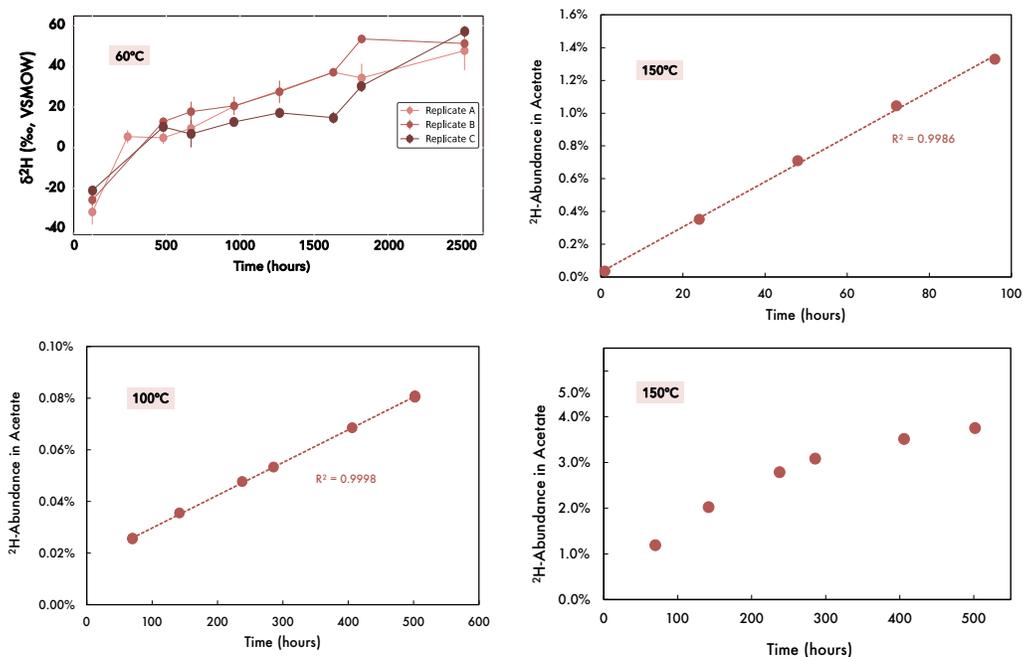
Figure 5.5: Exchange experiments in deuterated water (5% ^{2}H2O) with 1 mM acetate at 60°C, 100°C, and 150°C. Acetate incorporated deuterium from the water into its methyl-site in a linear fashion. At 60°C, the exchange rate is so low that deuterium incorporation is quantified as changes in natural abundance $\delta^2 H$ values (10‰ is equivalent to 1 ppm absolute increase). All measurements made on an electrospray Orbitrap mass spectrometer.
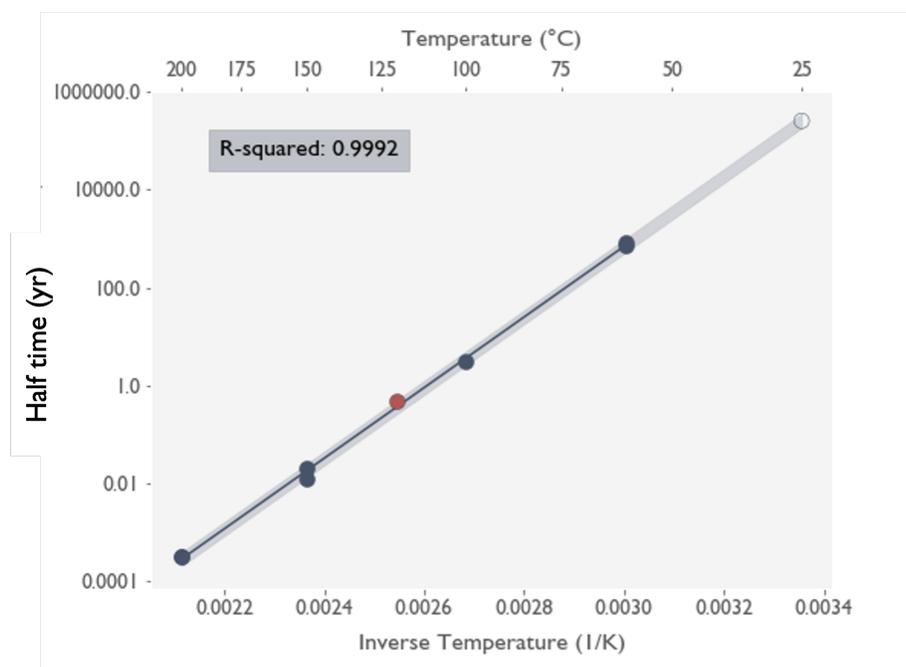
Figure 5.6: Arrhenius plot of exchange reaction rate with temperature (from Figure 5.1 in main text) for acetate in deuterated water (5% $2H_2O$). Dark blue dots represent exchange with pure water. Red dot represents exchange in 1M $CaCl_2$ at 120°C.
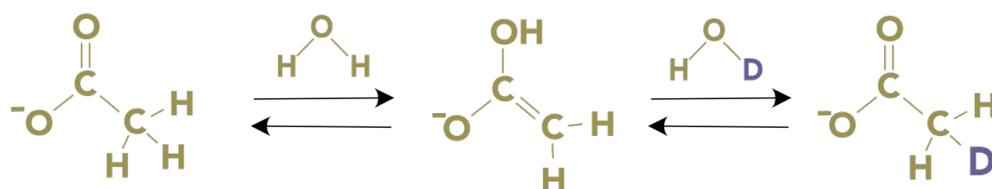
Figure 5.7: Proposed tautomerization reaction mechanism for exchanging hydrogen isotopes between water and the methyl group of acetate.

Table 5.3: Composition of synthetic solution used for validation studies on acetate purification techniques.

| Location | Kidd Creek | Kidd Creek | Kidd Creek | Kidd Creek | Kidd Creek | Birch Tree | Birch Tree | Birch Tree |
Borehole	12261	12299	12287	12299	12261	1113860	1113740	BH1A
Average $\delta^{13}C$	-8.8‰	-7.0‰	-6.6‰	-8.6‰	-10.0‰	-26.8‰	-27.5‰	-26.7‰
Error $\delta^{13}C$	0.5‰	0.2‰	0.8‰	1.2‰	0.9‰	0.9‰	0.7‰	0.6‰
Average δ^2H	-134‰	-130‰	-143‰	-134‰	-130‰	-170‰	-167‰	-170‰
Error δ^2H	3‰	4‰	2‰	3‰	4‰	5‰	7‰	8‰
Water δ^2H	-36‰	-36‰	-36‰	-36‰	-36‰	-74‰	-74‰	-74‰
Alpha (ac/w)	0.898	0.903	0.890	0.899	0.902	0.897	0.900	0.896
Acetate Conc. (mM)**	1.6	1.4	1.9	1.4	1.6	0.15	0.23	0.27
Temperature (°C)	25	21	21	21	25	21	21	21
Depth**	2.4 km	2.4 km	2.4 km	2.4 km	2.4 km	—	—	—
Sampling Date	28/08/2007	27/08/2007	28/08/2007	27/08/2007	2018	07/11/2007	08/11/2007	28/05/2007
Method	Dual Inlet	Dual Inlet	Dual Inlet	Dual Inlet	Dual Inlet	Direct Inf.	Direct Inf.	Direct Inf.

References

Beulig, F. et al. (Jan. 2018). "Control on rate and pathway of anaerobic organic carbon degradation in the seabed". en. In: *Proc. Natl. Acad. Sci. U.S.A.* 115.2, pp. 367–372. DOI: 10.1073/pnas.1715789115.

Blair, N.E. and W.D. Carter (Mar. 1992). "The carbon isotope biogeochemistry of acetate from a methanogenic marine sediment". en. In: *Geochimica et Cosmochimica Acta* 56.3, pp. 1247–1258. DOI: 10.1016/0016-7037(92)90060-V.

Blaser, Martin B., Lisa K. Dreisbach, and Ralf Conrad (Mar. 2013). "Carbon Isotope Fractionation of 11 Acetogenic Strains Grown on H $_2$ and CO $_2$". en. In: *Appl Environ Microbiol* 79.6, pp. 1787–1794. DOI: 10.1128/AEM.03203-12.

Bomberg, Malin et al. (2021). "Microbial metabolic potential in deep crystalline bedrock". en. In: *The Microbiology of Nuclear Waste Disposal*. Elsevier, pp. 41–70. DOI: 10.1016/B978-0-12-818695-4.00003-4.

Bradley, J. A. et al. (Aug. 2020). "Widespread energy limitation to life in global subseafloor sediments". en. In: *Sci. Adv.* 6.32, eaba0697. DOI: 10.1126/sciadv.aba0697.

Costagliola, Amaury et al. (Nov. 2017). "Radiolytic Dissolution of Calcite under Gamma and Helium Ion Irradiation". en. In: *J. Phys. Chem. C* 121.44, pp. 24548–24556. DOI: 10.1021/acs.jpcc.7b07299.

Etiope, Giuseppe and Barbara Sherwood Lollar (Apr. 2013). "ABIOTIC METHANE ON EARTH". en. In: *Reviews of Geophysics* 51.2, pp. 276–299. DOI: 10.1002/rog.20011.

Ferguson, Grant et al. (Aug. 2021). "Crustal Groundwater Volumes Greater Than Previously Thought". en. In: *Geophysical Research Letters* 48.16, e2021GL093549. DOI: 10.1029/2021GL093549.

Gao, Jiali (Oct. 1996). "A theoretical investigation of the enol content of acetic acid and the acetate ion in aqueous solution". en. In: *Journal of Molecular Structure: THEOCHEM* 370.2-3, pp. 203–208. DOI: 10.1016/S0166-1280(96)04702-1.

Gelwicks, J T, J B Risatti, and J M Hayes (Feb. 1994). "Carbon isotope effects associated with aceticlastic methanogenesis". en. In: *Appl Environ Microbiol* 60.2, pp. 467–472. DOI: 10.1128/aem.60.2.467-472.1994.

Gelwicks, Jeffrey T., J.Bruno Risatti, and J.M. Hayes (Jan. 1989). "Carbon isotope effects associated with autotrophic acetogenesis". en. In: *Organic Geochemistry* 14.4, pp. 441–446. DOI: 10.1016/0146-6380(89)90009-0.

Heuer, Verena et al. (Nov. 2006). "Online δ ^{13} C analysis of volatile fatty acids in sediment/porewater systems by liquid chromatography-isotope ratio mass spectrometry: Online δ ^{13} C analysis of VFAs". en. In: *Limnol. Oceanogr. Methods* 4.10, pp. 346–357. DOI: 10.4319/lom.2006.4.346.

Heuer, Verena B. et al. (June 2009). "The stable carbon isotope biogeochemistry of acetate and other dissolved carbon species in deep subseafloor sediments at the northern Cascadia Margin". en. In: *Geochimica et Cosmochimica Acta* 73.11, pp. 3323–3336. DOI: 10.1016/j.gca.2009.03.001.

Hilkert, Andreas et al. (July 2021). "Exploring the Potential of Electrospray-Orbitrap for Stable Isotope Analysis Using Nitrate as a Model". en. In: *Anal. Chem.* 93.26, pp. 9139–9148. DOI: 10.1021/acs.analchem.1c00944.

Holland, G. et al. (May 2013). "Deep fracture fluids isolated in the crust since the Precambrian era". en. In: *Nature* 497.7449, pp. 357–360. DOI: 10.1038/nature12127.

Jørgensen, Bo Barker, Alyssa J. Findlay, and André Pellerin (Apr. 2019). "The Biogeochemical Sulfur Cycle of Marine Sediments". en. In: *Front. Microbiol.* 10, p. 849. DOI: 10.3389/fmicb.2019.00849.

Lee, Chengteh, Weitao Yang, and Robert G. Parr (Jan. 1988). "Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density". en. In: *Phys. Rev. B* 37.2, pp. 785–789. DOI: 10.1103/PhysRevB.37.785.

Lever, Mark A. et al. (Feb. 2010). "Acetogenesis in Deep Subseafloor Sediments of The Juan de Fuca Ridge Flank: A Synthesis of Geochemical, Thermodynamic, and Gene-based Evidence". en. In: *Geomicrobiology Journal* 27.2, pp. 183–211. DOI: 10.1080/01490450903456681.

Li, L. et al. (Oct. 2016). "Sulfur mass-independent fractionation in subsurface fracture waters indicates a long-standing sulfur cycle in Precambrian rocks". en. In: *Nat Commun* 7.1, p. 13252. DOI: 10.1038/ncomms13252.

Lin, Li-Hung et al. (July 2005). "Radiolytic H$_2$ in continental crust: Nuclear power for deep subsurface microbial communities". en. In: *Geochem Geophys Geosyst* 6.7, 2004GC000907. DOI: 10.1029/2004GC000907.

Lin, Li-Hung et al. (Feb. 2005). "The yield and isotopic composition of radiolytic H2, a potential energy source for the deep subsurface biosphere". en. In: *Geochimica et Cosmochimica Acta* 69.4, pp. 893–903. DOI: 10.1016/j.gca.2004.07.032.

Lollar, Barbara Sherwood et al. (Dec. 2014). "The contribution of the Precambrian continental lithosphere to global H2 production". en. In: *Nature* 516.7531, pp. 379–382. DOI: 10.1038/nature14017.

Lollar, Garnet S. et al. (Nov. 2019). "'Follow the Water': Hydrogeochemical Constraints on Microbial Investigations 2.4 km Below Surface at the Kidd Creek Deep Fluid and Deep Life Observatory". en. In: *Geomicrobiology Journal* 36.10, pp. 859–872. DOI: 10.1080/01490451.2019.1641770.

Magnabosco, Cara et al. (Mar. 2016). "A metagenomic window into carbon metabolism at 3 km depth in Precambrian continental crust". en. In: *The ISME Journal* 10.3, pp. 730–741. DOI: 10.1038/ismej.2015.150.

Mardyukov, Artur, André K. Eckhardt, and Peter R. Schreiner (Mar. 2020). "1,1-Ethenediol: The Long Elusive Enol of Acetic Acid". en. In: *Angew Chem Int Ed* 59.14, pp. 5577–5580. DOI: 10.1002/anie.201915646.

Mueller, Elliott P., Alex L. Sessions, et al. (Jan. 2022). "Simultaneous, High-Precision Measurements of $\delta\ ^2$H and $\delta\ ^{13}$C in Nanomole Quantities of Acetate Using Electrospray Ionization-Quadrupole-Orbitrap Mass Spectrometry". en. In: *Anal. Chem.* 94.2, pp. 1092–1100. DOI: 10.1021/acs.analchem.1c04141.

Mueller, Elliott P., Fenfang Wu, and Alex L. Sessions (Nov. 2022). "Quantifying Isotopologue Reaction Networks (QIRN): A modelling tool for predicting stable isotope fractionations in complex networks". en. In: *Chemical Geology* 610, p. 121098. DOI: 10.1016/j.chemgeo.2022.121098.

Nisson, D.M. et al. (Jan. 2023). "Hydrogeochemical and isotopic signatures elucidate deep subsurface hypersaline brine formation through radiolysis driven water-rock interaction". en. In: *Geochimica et Cosmochimica Acta* 340, pp. 65–84. DOI: 10.1016/j.gca.2022.11.015.

Ouellette, Robert J. and J. David Rawn (2015). "Condensation Reactions of Carbonyl Compounds". en. In: *Organic Chemistry Study Guide*. Elsevier, pp. 419–463. DOI: 10.1016/B978-0-12-801889-7.00022-4.

Sessions, Alex L. et al. (Apr. 2004). "Isotopic exchange of carbon-bound hydrogen over geologic timescales 1 1Associate editor: J. Horita". en. In: *Geochimica et Cosmochimica Acta* 68.7, pp. 1545–1559. DOI: 10.1016/j.gca.2003.06.004.

Sherwood Lollar, B., S.K. Frape, et al. (Dec. 1993). "Abiogenic methanogenesis in crystalline rocks". en. In: *Geochimica et Cosmochimica Acta* 57.23-24, pp. 5087–5097. DOI: 10.1016/0016-7037(93)90610-9.

Sherwood Lollar, B., V.B. Heuer, et al. (Feb. 2021). "A window into the abiotic carbon cycle – Acetate and formate in fracture waters in 2.7 billion year-old host rocks of the Canadian Shield". en. In: *Geochimica et Cosmochimica Acta* 294, pp. 295–314. DOI: 10.1016/j.gca.2020.11.026.

Stephens, P. J. et al. (Nov. 1994). "Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields". en. In: *J. Phys. Chem.* 98.45, pp. 11623–11627. DOI: 10.1021/j100096a001.

Taguchi, Koudai et al. (Oct. 2022). "Low 13C-13C abundances in abiotic ethane". en. In: *Nat Commun* 13.1, p. 5790. DOI: 10.1038/s41467-022-33538-9.

Tappe, W. et al. (June 1999). "Maintenance Energy Demand and Starvation Recovery Dynamics of *Nitrosomonas europaea* and *Nitrobacter winogradskyi* Cultivated in a Retentostat with Complete Biomass Retention". en. In: *Appl Environ Microbiol* 65.6, pp. 2471–2477. DOI: 10.1128/AEM.65.6.2471-2477.1999.

Telling, Jon et al. (Feb. 2018). "Bioenergetic Constraints on Microbial Hydrogen Utilization in Precambrian Deep Crustal Fracture Fluids". en. In: *Geomicrobiology Journal* 35.2, pp. 108–119. DOI: 10.1080/01490451.2017.1333176.

Templeton, Alexis S. and Tristan A. Caro (May 2023). "The Rock-Hosted Biosphere". en. In: *Annu. Rev. Earth Planet. Sci.* 51.1, pp. 493–519. DOI: 10.1146/annurev-earth-031920-081957.

Thomas, Randal B (2008). "Intramolecular isotopic variation in acetate in sediments and wetland soils." en. In.

Tijhuis, L., M. C. M. Van Loosdrecht, and J. J. Heijnen (Aug. 1993). "A thermodynamically based correlation for maintenance gibbs energy requirements in aerobic and anaerobic chemotrophic growth". en. In: *Biotech & Bioengineering* 42.4, pp. 509–519. DOI: 10.1002/bit.260420415.

Urey, Harold C. (1947). "The thermodynamic properties of isotopic substances". en. In: *J. Chem. Soc.*, p. 562. DOI: 10.1039/jr9470000562.

Van Bodegom, Peter (May 2007). "Microbial Maintenance: A Critical Review on Its Quantification". en. In: *Microb Ecol* 53.4, pp. 513–523. DOI: 10.1007/s00248-006-9049-5.

Vandenborre, Johan et al. (June 2021). "Carboxylate anion generation in aqueous solution from carbonate radiolysis, a potential route for abiotic organic acid synthesis on Earth and beyond". en. In: *Earth and Planetary Science Letters* 564, p. 116892. DOI: 10.1016/j.epsl.2021.116892.

Warr, Oliver, Thomas Giunta, et al. (Feb. 2021). "The role of low-temperature 18O exchange in the isotopic evolution of deep subsurface fluids". en. In: *Chemical Geology* 561, p. 120027. DOI: 10.1016/j.chemgeo.2020.120027.

Warr, Oliver, Min Song, and Barbara Sherwood Lollar (Mar. 2023). "The application of Monte Carlo modelling to quantify in situ hydrogen and associated element production in the deep subsurface". en. In: *Front. Earth Sci.* 11, p. 1150740. DOI: 10.3389/feart.2023.1150740.

Wellmer, F.-W. et al. (Jan. 1999). "Carbon Isotope Geochemistry of Archean Carbonaceous Horizons in the Timmins Area". In: DOI: 10.5382/Mono.10.18.

Chapter 6

SHIFTS IN RUMINANT FERMENTATION PATHWAYS DURING METHANOGENIC INHIBITION ARE RECORDED IN THE ISOTOPE COMPOSITION OF RUMEN ORGANIC ACIDS.

Elliott P. Mueller[1*], Rich Duong[2], John Eiler[1], Matthias Hess[2], Alex L. Sessions [1]

[1]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA [2]College of Agriculture and Environmental Science, University of California Davis, Davis, CA, USA
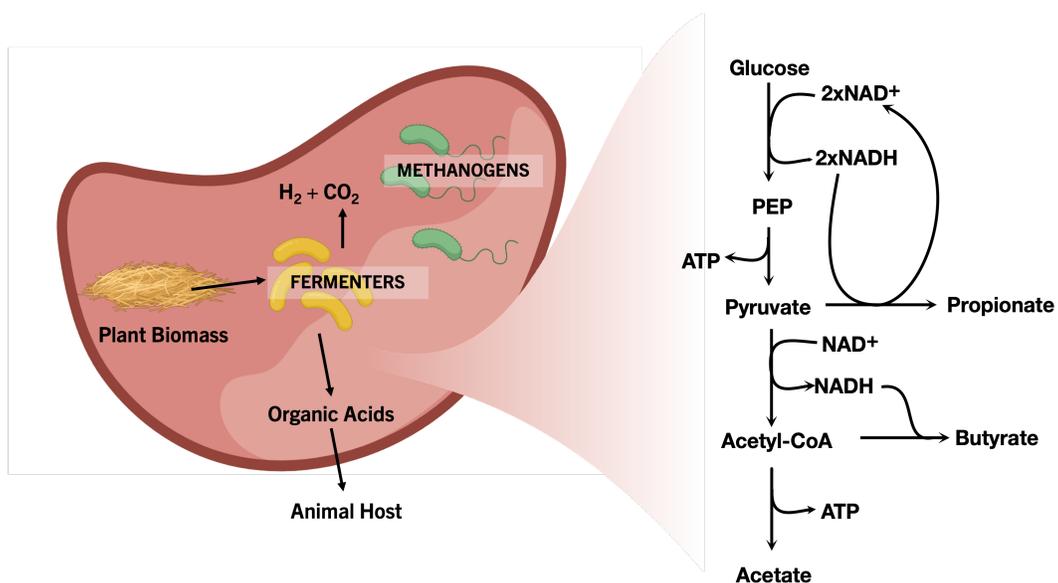
Graphical Abstract of Chapter 6

6.1 Abstract

Ruminant animals (e.g. cows) are a major source of the potent greenhouse gas methane, but they are also a tractable target for climate solutions. Multiple strategies have been developed to lower methane emissions from ruminants, including feed additives that inhibit the resident population of methanogenic microbes. However, sustainable strategies must elimate methane emissions without hampering the microbial fermentation of plant material. Fermenting bacteria generate volatile fatty acids (VFAs) for the animal host to use as its primary source of carbon and energy. As such, fermentation is an essential process to maintain in the rumen. Typically, fermenting organisms live in syntrophic relationships with methanogenic microbes that consume the waste products of fermentation, CO_2 and H_2 gas. Supplementing feed with methanogen-inhibiting additives will break this syntrophy, and fermenting bacteria will respond. However, we currently lack the tools to quantify how carbon flow through fermentation is altered under these new conditions. Here, we developed a new method to quantify fermentative pathways in the rumen using the stable isotope ratios ($^{13}C/^{12}C$ and $^{2}H/^{1}H$) of VFAs at their natural abundances. We tested this technique on incubations of rumen fluid with and without the additive *Asparagopsis taxiforms*, a red macroalgae that inhibits methanogens. We found that the isotope composition of VFAs changed between these conditions, reflecting a metabolic shift in the fermentation of plant material. Our results indicate that an increase in acetate's ^{13}C isotope ratio corresponds to a decrease in its synthesis. Meanwhile, the δ^2H value of propionate decreased with time only when *A. taxiforms* was added to incubations, likely reflecting a shift between the two pathways of propionate synthesis. Thus, the δ^2H and $\delta^{13}C$ values of VFAs in the rumen indicate that acetate synthesis slows down and propionate synthesis changes pathways when methanogenesis is inhibited. Our study demonstrates that the stable isotope composition of VFAs will be a useful tool for evaluating attempts to direct H_2 flow away from methanogenesis and toward alternative H_2 sinks like reductive acetogenesis and propionic acid fermentation. ESI-Orbitrap-based isotope ratio analysis of VFAs could be a useful method for evaluating the efficacy and sustainability of engineered methane mitigating strategies in bovine rumen.

6.2 Introduction

Methane is a potent greenhouse gas responsible for more than 20% of global warming (Etminan et al., 2016). Though lower in concentration than CO_2 , methane traps solar radiation in the Earth's atmosphere 25-times more efficiently, amplifying its

contributions to climate change (Smith et al., 2011). Fortunately, methane has a short (decadal) residence time in the atmosphere. Thus, climate solutions that mitigate methane emissions could have meaningful impacts within a human lifetime.

Livestock agriculture, particularly of ruminant animals (e.g. cows), is the second largest source of anthropogenic methane behind fossil fuels and a tractable target for such mitigation efforts (National Academies of Sciences and Medicine, 2018; Patra et al., 2017). The anaerobic degradation of fibrous plant materials in the cow rumen creates nearly 500 liters of methane per day per animal (K. A. Johnson and D. E. Johnson, 1995). With an estimated 1.5 billion cows on the planet, enteric fermentation is responsible for at least 6% of climate change (National Academies of Sciences and Medicine, 2018). In recent years, multiple strategies have been developed to lower methane emissions from cows, including feed-additives. Dosed at low concentrations (<2% w/w), these additives have been shown to drastically decrease methane production in ruminant animals (Patra et al., 2017; Roque et al., 2019). In order to test the sustainability of these solutions, we require an understanding of the underlying mechanism that gave them preliminary success and the consequences of their use on the fermentation process.

Methane is made in the cow rumen through a cascade of metabolic reactions that breaks down solid plant matter (Czerkawski, 1986). These processes are driven by microorganisms — bacteria, archaea, fungi and protists — densely populating the rumen ($\sim 10^{11}$ cells/mL) (Nagaraja, 2016). While the physical mechanisms of ruminant fermentation are still an area of active research, the overall structure of carbon flow is well defined. Cellulose — a biopolymer of glucose monomers bound by heard-to-break beta linkages — is first cleaved by exoenzymes excreted from microbial cells. The released glucose is then broken down by any number of fermentative metabolic pathways. Rather than relying on a terminal electron acceptor as an oxidant, fermenting microbes use glucose as both an electron donor and acceptor creating CO_2 and volatile fatty acids (VFAs), respectively (White, Drummond, and Fuqua, 2012). The animal host absorbs these VFAs and consume them for energy (Nagaraja, 2016). Cultivation of fermentative microorganisms in the rumen allows ruminant animals to convert energy-dense plant matter into bioavailable carbon. However, many fermenting microbes cannot maintain redox balance with VFAs as the only sink of reducing equivalents. They also use water as an electron acceptor, which generates hydrogen gas (H_2). Methanogenic archaea consume this H_2 and CO_2 to make methane. About 12% of the carbon consumed

by the bovine host is ultimately released as methane (Beauchemin and McGinn, 2006). Thus, enteric methane production is not only a significant contributer to anthropogenic climate change, but also an energetic loss for the host. Strategies to eliminate methanogenic archaea from the gut microbiome have become a subject of ongoing research. Many of these strategies seek to take one step further, replacing methanogens with microbial H_2 sinks that divert CO_2 to further VFA production (e.g. acetogenesis) (E. M. Ungerfeld, 2015; E. Ungerfeld, 2020).

Feed additives have recently gained traction as an option to mitigate methane emissions from cow rumen fermentation (Patra et al., 2017; Roque et al., 2019; E. Ungerfeld, 2020; Kung et al., 2003). These additives are typically methanogen-specific metabolic inhibitors, namely small halogenated compounds (E. M. Ungerfeld, 2015). They can be added directly to the feed as the purified bioactive compound or within plant materials that naturally accumulate the compounds in their tissue. During the fermentation process, they are released into the rumen fluid and bind to the active site of methyl-coenzyme M reductase (MCR), a requisite enzyme for methanogenesis (Romero et al., 2023). By inactivating MCR, these compounds severely limit methane emissions during *in vitro* incubations of rumen fluid (Glasson et al., 2022). They have also been shown to work *in vivo* (i.e. in the animal host). The macroalgae *Asparagopsis taxiforms* naturally contains large concentrations of bromoform, an effective MCR inhibitor. When added to feeds it can lower methane production rates by up to 99% (Brooke et al., 2020; Machado et al., 2014; Maia et al., 2016). It can also be grown in a variety of coastal environments. If enough is produced through marine farming, *A. taxiforms* could be a sustainable mechanism of inhibiting methane emissions in cows globally.

While additives successfully lower methane production, eliminating a major metabolic pathway in the rumen system has consequences for the residual community of fermenting microorganisms. Fermentation and methanogenesis form an syntrophy that expedites plant matter degradation (Hobson and Stewart, 1997). Fermenters feed methanogens with H_2 and CO_2 while methanogens continually remove the waste products of fermentation, enabling rapid cellulose decomposition that sustains the animal host. If H_2 accumulates — like when methanogens are inhibited — fermentation gains less free energy per mole of glucose consumed, slowing down the total rate of organic degradation (E. Ungerfeld, 2020). Thus, eliminating methanogenic communities has profound implications for the rumen's microbial ecology. Indeed, *in vitro* incubations of rumen fluid with *A. taxiforms* have lower VFA production

rates than controls with the same amount of feed (Roque et al., 2019). The relative abundance of acetate, propionate and butyrate - the major VFAs in the rumen - also changes during methanogen inhibition (E. M. Ungerfeld, 2015; Kung et al., 2003). Since these changes are a consequence of the decoupled syntrophy between methanogens and fermenters, they will occur irrespective of the methane mitigation strategy invoked. Even if microbial communities can be engineered to exclude methanogens during the initial colonization of the rumen, thermodynamic barriers within the cascade of metabolic pathways that breaks down cellulose will still exist. Understanding the response of fermenting communities both in structure and in function is crucial for evaluating the efficacy of engineered solutions and the sustainability of their implementation.

Sustainable strategies will eliminate methane emissions without hampering the fermentation of plant material; however, this is difficult to evaluate in practice. Since multiple fermentative pathways can create the same VFA products, identifying how these pathways change when methanogens are inhibited is challenging even during *in vitro* incubations. The most common metrics are the total VFA production rate and the relative abundance of these VFAs (E. M. Ungerfeld, 2015; Kung et al., 2003). These can shed light on the response of fermentative organisms; however, their application will become even more limited when strategies are implemented *in vivo*. In the rumen, organic acids are held in a steady state by the animal host continually consuming VFAs as they are produced (Czerkawski, 1986). In this case, profiles of VFA concentrations will be less informative and their total production rate more difficult to measure. Without isotopically labeled substrates, which are too expensive to scale-up, identifying the fermentative pathways creating these VFAs has been impossible. And yet the mechanisms and rates of enteric fermentation are essential for the health and growth of the animal host. We must develop tools that test the response of fermentation to methane mitigating strategies and translate to *in vivo* studies

Engineered solutions further attempt to modify the rumen carbon flow by diverting H_2 to pathways other than methanogenesis, including to propionate production or reductive acetogenesis. In both cases, these alternative metabolisms replace methanogenesis as the syntrophic partner to fermentation, spurring plant degradation while also producing additional VFAs, an energetic boon for the animal host (E. M. Ungerfeld, 2015). Since these pathways create the same molecules as the aforementioned fermentative metabolisms, their successful integration into the com-

munity can only be evaluated with metagenomic analyses. Genetic tools can identify a sustained population of organisms that use H_2 for acetogenesis or propionic acid production, but they do not give information about actual metabolic utilization by these taxa or their production rate (Greening et al., 2019). Here, we developed a novel, high-throughput analytical tool that measures the stable isotope composition of VFAs at their natural abundance in the cow rumen. We demonstrate that this additional layer of information on top of VFA profiling can yield useful information about fermentation mechanisms and the presence of alternative H_2-sinks. They are also more likely to be translatable to *in vivo* studies.

Stable isotopes at their natural abundances are a useful tool for understanding chemical and biological processes in nature. These tools capitalize on the natural variability of isotopes like ^{13}C and ^2H in biomolecules. The addition of an extra neutron to the atom does not change its broad chemical properties; however, enzymatic reactions often express a kinetic preference for those molecules with lighter (^{12}C) isotopes over those containing heavier (^{13}C) isotopes. This kinetic isotope effect (KIE) or 'fractionation' is realized as a measurable difference between the isotope ratio ($^{13}C/^{12}C$) of the reaction's substrate and its product. The isotope compositions (i.e. $\delta^{13}C$ and δ^2H) of molecules are expressed as a perthousand or 'permil' (‰) differences in isotope ratio relative to an internationally recognized standard material.

Since the magnitude of an enzyme's KIE is dependent on its mechanism, metabolic pathways with different enzymes will express distinct KIEs (J. M. Hayes, 2001). By measuring the $\delta^{13}C$ or δ^2H composition of molecules in nature, we can distinguish metabolic sources of the same molecules. For example, fermentation and acetogenesis both create acetate, the most common VFA in rumen fluid. However, the $\delta^{13}C$ or δ^2H composition of acetate is different depending on which metabolism produced it. Fermentation tends to produce acetate with an enriched $\delta^{13}C$ and δ^2H compositions (0-8‰ enriched relative to the substrate). Whereas acetogenesis expresses a strong (>40‰) isotope effect on the CO_2 used as a substrate, creating acetate with a strongly ^{13}C-depleted signature (Mueller et al., 2022; Penning and Conrad, 2006; Gelwicks, Risatti, and J. Hayes, 1989; Blaser, Dreisbach, and Conrad, 2013). As demonstrated in Chapter 4, the exact fermentation pathway utilized to produce acetate and other VFAs also effects their isotope compositions.

We hypothesized that the isotope compositions of VFAs in the rumen could be a useful constraint on the changing pathways of fermentation upon addition of *A.*

taxiforms. However, to our knowledge, they had never been measured in cow rumen before. We used a novel, electrospray ionization (ESI) Orbitrap mass spectrometry (MS) technique for high-throughput analysis of acetate's $\delta^{13}C$ and δ^2H values (Mueller et al., 2022). We further expanded this technique to simultaneously characterize the isotope compositions of propionate and butyrate. For the analysis of rumen fluid, this method required less than a drop of sample, requires almost no sample preparation, and makes high precision measurements in less than twenty minutes. To test our hypothesis, we collected fluid and emitted gases from three-day incubations of rumen fluid fed a series of organic substrates with and without *A. taxiforms*. In addition to the classic decrease in methane emissions and VFA production rates, we see reproducible shifts in the isotope compositions of acetate and propionate. These signals indicate that acetate is either synthesized at much lower rates or converted to other organic acids when methanogens are inhibited and that acetogenic bacteria are not making significant amounts of acetate under any condition. Our results may also suggest that propionate production shifts from the succinate pathway to the lactate pathway, based on a consistent depletion in the hydrogen isotope composition of propionate upon addition of *A. taxiforms*. Together, our study demonstrates that isotopic analysis of VFAs is a useful tool for identifying fermentation pathways in the rumen that are otherwise invisible. We predict that it could also quantify alternative H_2 pathways like propionate formation and acetogenesis in the rumen. Furthermore, these signals are likely to translate to *in vivo* experiments. Thus, ESI-Orbitrap-based isotope ratio analysis of VFAs could be a high-throughput method for evaluating the efficacy and sustainability of engineered methane mitigating strategies in bovine rumen.

6.3   Methods

Rumen Fluid Collection

All animal procedures were performed in accordance with the Institution of Animal Care and Use Committee (IACUC) at the University of California, Davis, under protocol number 19263. Rumen content was collected from a fistulated Holstein cow that housed at the UC Davis Dairy Research and Teaching Facility Unit. The donor had been fed TMR. Two liters of rumen fluid and 30 g of rumen solids were collected 90 min after morning feeding. Rumen content was collected via transphonation using a perforated PVC pipe, 500 mL syringe, and Tygon tubing (Saint-Gobain North America, PA, United States). Fluid was strained through a colander into a pre-warmed, vacuum insulated container and transported to the
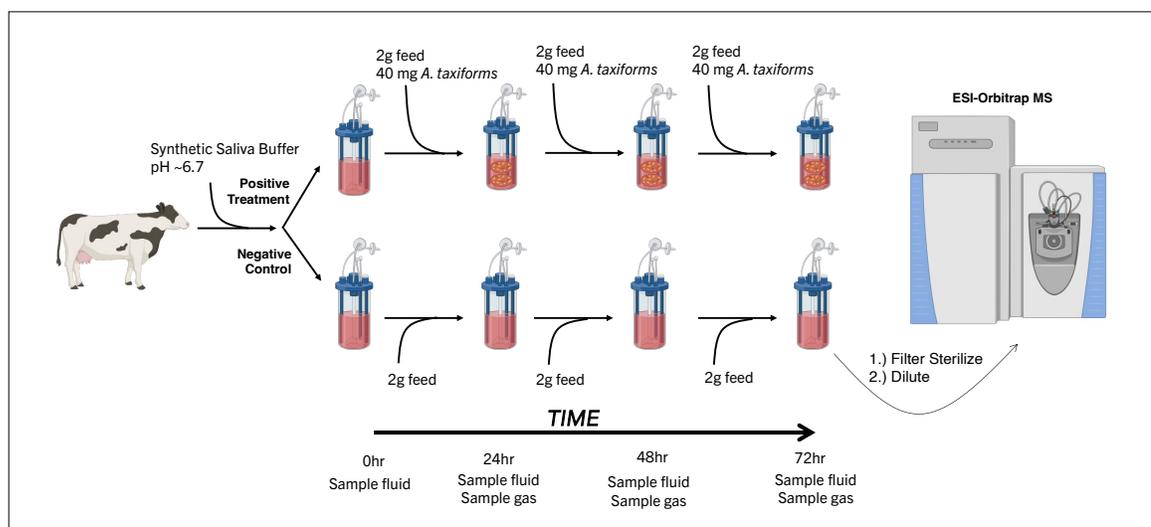
Figure 6.1: Overview of *in vitro* rumen incubation experiments. Rumen fluid was incubated with and without *A. taxiforms* additive in Ankom units for three days. Fluid and gas samples were taken every 24 hours. Organic acids in the fluid were diluted into methanol and then directly infused into an ESI-Orbitrap MS.

laboratory. Rumen fluid and solids were collected on November 13, 2023. Within 2 hours of collection, the trials had started.

Rumen Fluid Incubation and Sampling

Rumen incubations were performed in 0.3 L Ankom units (Ankom Technology RF Gas Production System, Macedon, NY, United States). Each unit received 200 mL of a 3:1 mixture of synthetic saliva buffer (Table 6.2) and rumen fluid. In addition, 2 grams of rumen solids and 2 grams of total mean ration (TMR) were added to each unit at the start of the trials. The composition of TMR (70% alfalfa, 15% dried distillers' grain and 15% rolled corn) and synthetic saliva buffer. Rumen solids were placed directly in the incubation while TMR was sealed in porous 5 cm × 5 cm concentrate feed bags. For alfalfa and cellulose treatments, 2 grams and 1.5 grams of feed was added, respectively, to the feed bags. *A. taxiforms* was included in the respective feed bags (Ankom, Macedon, NY, United States) at 2% (w/w). The Ankom units were placed and incubated in a shaking water bath (39°C, 40 rpm). Foil gas bags (Restek, United States) were filled with 30 mL of pure nitrogen gas and then connected to the Ankom units. With three different feeds, positive treaments with *A. taxiforms* , and negative controls without *A. taxiforms* , there were six conditions. Every condition included four Ankom replicates resulting in 24 total units run in parallel.

Throughout the experiments, the Ankom units automatically opened their valve between the gas bag and the incubation headspace when the headspace reached a set pressure. At 24 and 48 hours, each gas bag was replaced with a new one pre-filled with 30 mL of nitrogen gas. At the same time, the Ankom units were rapidly opened, the old feed bag removed, and a fresh one placed in the incubation. Their headspace were subsequently flushed with nitrogen gas for 30 seconds before closing the units and attaching the new gas bag. When the Ankom units were open, 1 mL of liquid was taken and immediately filter sterilized and placed at -20°C for organic acid analysis.

CO_2 and methane concentration measurements

CO_2 and methane were measured from gas bags every 24 hours using an SRI Gas Chromatograph (8610C, SRI, Torrance, CA, United States) fitted with a 3' × 1/8" stainless steel Haysep D column. The oven temperature was held at 90°C for 5 min. Carrier gas was high purity nitrogen at a flow rate of 30 mL/min. A 1 mL sample was diluted in 29 mL of pure nitrogen and injected directly onto the column. Calibration curves were developed with Airgas certified methane and CO_2 standard (Airgas, United States).

CO_2 and methane isotopic analysis

The carbon isotope composition of CO_2 and methane were measured on a gas chromatography isotope ratio mass spectrometer (GC-IRMS). A Thermo Scientific Trace 1310 gas chromatograph with a GC Isolink II was coupled to a 253 plus 10 kV IRMS. Injection volumes varied from 10 to 100 uL of gas samples through a 100 μL gas-tight syringe. The inlet temperature was held at 30°C. The injection volume and injector inlet split flow values were changed to match the heights of CO_2 and methane sample peaks with the CO_2 reference gas peak height of 7 volts. Typically, CO_2 concentrations were several-fold higher than methane concentrations. A split ratio of 40 was used for CO_2 measurements, while a split flow of 3 was used for methane, meaning the sample was injected on two separate occasions to measure each of the two analytes. Carrier gas flow rates were held at 2 mL/min. Using an Agilent PORAPLOT-Q 25 meter column, CO_2 and methane were sufficiently separated in a 6 minute, isothermal GC run held at 30°C. The $\delta^{13}C$ value of both analytes was anchored to the VPDB scale using a tank of pure CO_2 with known carbon isotope composition (-12.04‰). To verify the accuracy of the measurement, purified methane of known carbon isotope composition (-42.9‰, (Stolper et al.,

2014)) diluted to similar concentrations as the samples (1% for split ratio 3 and 10% for split ratio 40). This standard was run every three samples to monitor instrument performance. The measured and known values did not deviate by more than 1‰ throughout the instrument session. Since biological quadruplicates were available, analytical replicates for each sample were not run. Instead biological reproducibility served as the uncertainty, which is represented in the box plot of Figures 6.2 and 6.3.

VFA concentration measurements

VFA concentrations in rumen samples were measured using a Hewlett Packard Series 1100 high performance liquid chromatograph (HPLC) coupled to refractive index detector (RID). VFAs were separated on an Aminex HPX-87H (300 x 7.8 mm) column with an isocratic 8 mN sulfuric acid mobile phase. Peak areas were converted to concentrations with an external calibration curve of acetate, propionate, and butyrate from 0.1 to 50 mM. Before injection, all rumen samples were filter-sterilized and diluted 10-fold in 8 mN sulfuric acid. Injection volumes for sampels and standards were 10 μL. Blanks of Milli-Q water were run every 3-5 samples to monitor sample carryover, which was not measurable. Calibration curves were rerun at the beginning of every day of the instrument session.

VFA isotopic measurements

The compound-specific isotope composition ($\delta^2 H$ and $\delta^{13}C$) of acetate, propionate and butyrate were measured on a heated electrospray ionization (HESI) Orbitrap mass spectrometer (MS) (Thermo Scientific) attached to a Vanquish HPLC (Thermo Scientific) with a split-sampler and 100 μL sample loop. No LC column was used for these analyses. Each injection, 100 μL of sample was pulled through the sample loop and then injected into the 5 μL/min flow of LC-MS grade methanol (Fisher chemical, Optima), which carried the sample directly into the Orbitrap MS. After 19 minutes of measurements (2 minutes of dead volume followed by 17 minutes of sample analysis), the flow rate was increased to 35μL/min for two minutes to flush out the remaining sample and then decreased to 5μL/min again to prepare for the next sample. The ESI ionization parameters can be found in Table 6.1. This 19 minutes was split up into three segments, each focusing the quadrupole and Orbitrap on one of the three organic acids (Table 6.1). The raw mass spectra were converted to ion counts using the Makarov equation. The ion intensities from mass spectra were converted to counts (N) by applying an empirical factor (C_N = 4.4) derived by

Makarov and Denisov (2009) and the charge of the ion (z). (Makarov and Denisov, 2009; Eiler et al., 2017).

$$N = \left(\frac{S}{N_P}\right)\left(\frac{C_N}{z}\right)\sqrt{\frac{R_N}{R}} \tag{6.1}$$

As in Mueller et al., 2022, the isotope ratios of each organic acid were then calculated from the ion counts and further converted to the delta notation on the VPDB and VSMOW scales. Only data between 3 and 20 minutes were considered for isotope ratio analyses, as this represents the period during which the ion current was stable and in a plateau. On either side of this period, the ion current was either decreasing as sample infused into the instrument or was dropping as the last of the sample entered the mass spectrometer. Futhermore, only the periods designated for a given organic acid were used to calculate that organic acid's isotope composition (e.g. Acetate: 2 to 11 minutes; propionate: 11-17 minutes, butyrate: 17-20 minutes). Each biological replicate was run once and was bracketed by a standard that matched each organic acid's ion current intensity within 30%.

Propionate and Butyrate Standards

To anchor the measured ^{13}C and ^{2}H isotope ratios on the international reference frames Vienna Pee Dee Belemite (VPDB) and Vienna Standard Mean Ocean Water (VSMOW), standards of known $\delta^{13}C$ and $\delta^{2}H$ values were required. Sodium propionate (>99%) and sodium butyrate (>99%) was obtained from commercial sources to create this standards. Its synthetic or biological origins are unknown. Stock solutions (2M) were made up in deionized water (Milli-Q). Batches of working 2M standard stock solutions in Milli-Q water at were flash frozen in liquid nitrogen. To ensure homogeneity, stocks were kept as frozen aqueous solutions. Aliquots from these stocks were taken by thawing them at room temperature, inverting the vials to homogenize, aliquoting and immediately re-freezing the stocks. Each standard was measured via Elementary Analyzer (EA) coupled to an Isotope Ratio Mass Spectrometer (IRMS) to independently determine their isotopic compositions relative to the VSMOW and VPDB reference materials. Carbon isotope compositions were measured with combustion EA-IRMS. Samples were calibrated against glycine (-45.7‰) and Urea (-27.8‰) standard. Hydrogen isotope compositions were measured with thermal conversion EA-IRMS using an elemental chromium catalyst as discussed by Gehre et al., 2015; Mueller et al., 2022. Samples were calibrated to the VSMOW scale by analyzing USGS77 (polyethylene powder) and C36 n-alkane

Table 6.1: HESI-Orbitrap MS Parameters

ESI Parameters	
Polarity	Negative
Sheath Gas Flow	10
Auxiliary Gas Flow	3
Sweep Gas Flow	3
Spray Voltage	3.1
Spray Current	<0.2 μA
Auxiliary Gas Temperature	100°C
Capillary Temperature	320°C
Acetate MS Parameters	
Time	0 — 11 minutes
Quadrupole Filter Range	57 - 62 m/z
Resolution	60,000 (at 200 m/z)
AGC	1×10^6
Microscans	1
S-Lens Radio Frequency Level	60%
Propionate MS Parameters	
Time	11 — 17 minutes
Quadrupole Filter Range	72 - 75 m/z
Resolution	60,000 (at 200 m/z)
AGC	1×10^6
Microscans	1
S-Lens Radio Frequency Level	60%
Butyrate MS Parameters	
Time	17 — 24 minutes
Quadrupole Filter Range	85 - 90 m/z
Resolution	120,000 (at 200 m/z)
AGC	1×10^6
Microscans	1
S-Lens Radio Frequency Level	60%

#2 provided by Arndt Schimmelman (Indiana University). The sodium propionate and sodium butyrate standards are referred to as ProA and ButA. ProA and ButA had $\delta^2 H$ values of -110‰ (± 1.5‰) and -113‰ (± 1‰), respectively. Their $\delta^{13}C$ values were -34.3‰ (± 0.1‰) and -30.4‰ (± 0.1‰), respectively.

Feed bulk isotopic analysis

The carbon isotope composition of TMR, alfalfa and cellulose were measured using an Thermo Scientific elemental analyzer (EA) with a ConFlo system coupled to a Delta V IRMS. A total of 5 μg of carbon was weighed out into tin capsules. The capsules and samples were combusted together to create CO_2 gas which was separated from other gases on a GC column and then directed into the MS with a helium flow. The CO_2 ion peak height was similar to the CO_2 reference gas (-12.04‰). Carbon isotope compositions were corrected against external standards of glucose and glycine with known $\delta^{13}C$ values of -11 and -45.7‰, respectively.

6.4   Results

Gas Production and Isotope Composition

Methane

Cow rumen was incubated with feed bags containing three different organic substrates: total mean ration (TMR), alfalfa (ALF), and cellulose (CEL). Over the course of three days, the total methane production rates (mL/g feed) in TMR and ALF controls that did not contain *A. taxiforms* were similar. They decreased from 15 mL/g to 2-5 mL/g. With CEL, the decrease also occurred, though the total methane production rates were lower, starting only at 10 mL/g and decreasing to near zero (Figure 6.2). As a purely fibrous material with no easily digestible proteins or starches, CEL condition was expected to have lower overall gas production. The $\delta^{13}C$ of methane did not vary at the 24 hr time point, even with the addition of *A. taxiforms*. However, when fed ALF and CEL, methane's $\delta^{13}C$ value shifted to more depleted values by 5‰ over 72 hours. This was only observed in negative controls. When *A. taxiforms* was added the incubations, methane was too low to measure for concentration or isotope composition after 24 hours (Figure 6.2).

Carbon dioxide

Carbon dioxide production rates were also lower in CEL relative to TMR and ALF. However, the trend of decreasing CO_2 production over time was not as clear in TMR and ALF as in the methane data (Figure 2). For every substrate, CO_2 production rates between incubations with and without *A. taxiforms* were similar. We note that CO_2 concentrations in the gas bags collected at 24 hour and 48 hour timepoints exceeded the operational range of the instrument. More clear was the temporal trend in carbon isotopes for CO_2 . Over the three days, the $\delta^{13}C$ of CO_2 decreased in every incubation with TMR or ALF. The temporal change was 5-10‰; however, the positive and negative treatments were also offset from one another. In TMR
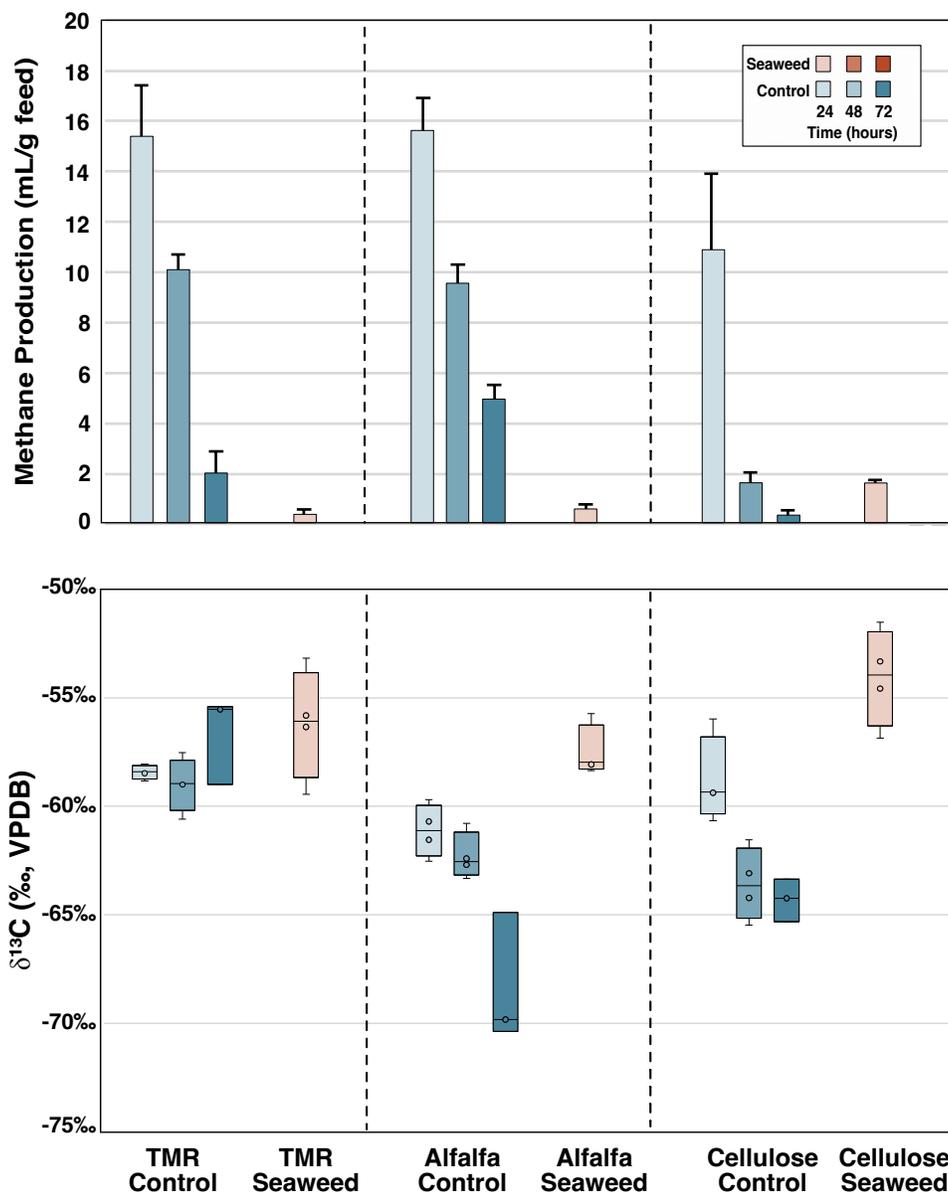
Figure 6.2: Methane emission rates in the rumen incubations and methane carbon isotope composition over three days. Methane emissions decreased by >90% with the addition of *A. taxiforms* . Blue bars and boxes indicate negative controls without *A. taxiforms* while red boxes indicate a position treatment with *A. taxiforms* .
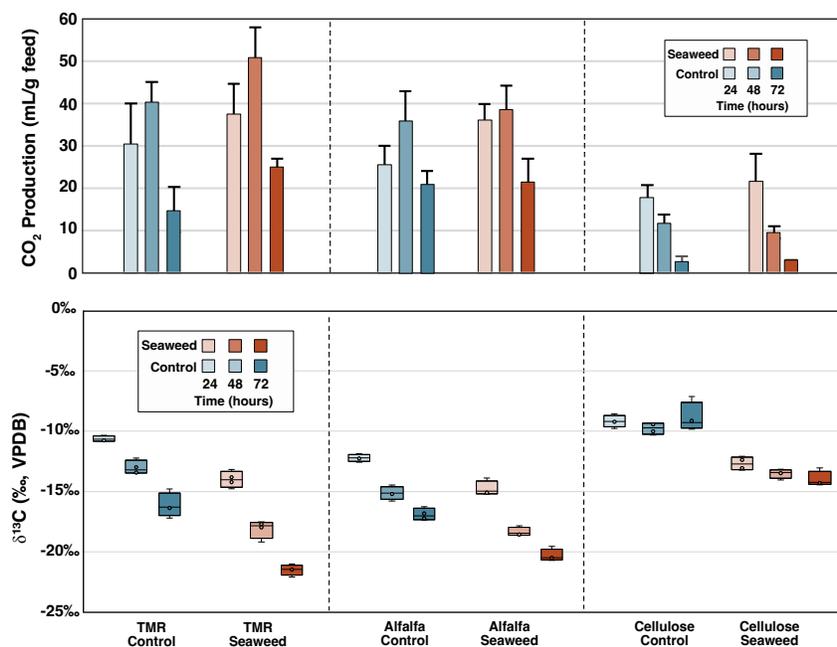
Figure 6.3: Carbon dioixed production rates in the rumen incubations and its carbon isotope composition over three days. Carbon dioxide production increased slightly upon addition of *A. taxiforms* , while its $\delta^{13}C$ balue decreased by ~5‰ between incubations with (red) and without (blue) *A. taxiforms* . The CO_2 $\delta^{13}C$ values decreased with time when ALF and TMR were used as feed, but not CEL.

and ALF incubations, those incubations with *A. taxiforms* was consistently 2-5‰ more depleted than the same time point in negative controls (Figure 2). When the incubations were fed CEL, there was no temporal change in $\delta^{13}C$ values of CO_2 , but the isotope compositions were offset by 2-5‰ between incubations with and without *A. taxiforms*. CO_2 produced from those incubations with methane inhibited were more ^{13}C-depleted, presumably due to the lack of a consumption reaction which would otherwise leave the residual pool of CO_2 ^{13}C-enriched.

VFA Production and Isotope Composition

The total production of VFAs decreased when methanogenesis was inhibited by the addition of *A. taxiforms* . When fed TMR, ALF, and CEL, negative controls produced 64, 77, and 38 mM VFAs respectively, combining acetate, propionate and butyrate. In positive treatments with *A. taxiforms* , the production for TMR, ALF, and CEL were 39, 57, and 25 mM, respectively. Error bars in Figure 6.3 indicate the standard error of the mean (SEM) of quadruplicate Ankom units. Overall, the SEM was less than 10% of the mean for each VFA. CEL had the least change in

VFA production with acetate being the only VFA that changed outside of error. This suggested that fermentation did not slow down as much as other conditions when methanogenesis was inhibited. The total VFA production was highest when ALF was added to the incubations. Since TMR contains more fibrous materials than pure ALF, the ratio of easily degraded starches and proteins to fibers in the ALF would have been higher than TMR. This also alligns with CEL having the slowest fermentation rates since it is entirely fibrous. Notably this trend is largely driven by acetate, the most abundant VFA in every incubation tested here.

The carbon isotope compositions of the VFAs had a consistent inverse relationship with carbon chain length. Acetate (C_2), propionate (C_3), and butyrate (C_4) had successively more depleted $\delta^{13}C$ values in the initial rumen fluid. These trends were less consistent as the incubations proceeded, though the overall pattern of $\delta^{13}C$ $_{Acetate}$ >$\delta^{13}C$ $_{Propionate}$ >$\delta^{13}C$ $_{Butyrate}$ remained. The hydrogen isotope composition of the three organic acids were remarkably similar, all around -210‰. In negative controls, this held true throughout the experiments, but this was not the case when *A. taxiforms* was added to the incubations. Overall, the biological reproducibility of the carbon and hydrogen isotope ratios was notably high with all four incubation replicates falling within analytical error (0.5‰ for carbon, 5‰ for hydrogen) under every condition. Error bars in Figure 6.4 represent the SEM of these replicates and are mostly encapsulated by the data markers. This level of reproducibility suggests that the isotope compositions of the VFAs are a robust signature of carbon flow. Below, we highlight key results from each VFA in terms of concentration and isotope composition.

Acetate

Acetate concentrations changed with time, the type of feed, and the presence or absence of methane inhibitory seaweed. In all negative controls (no macroalgae), acetate concentrations rose monotonically from 17mM to 50 mM, 65 mM, and 40 mM with TMR, ALF, and CEL respectively. However, when *A. taxiforms* was added to the incubations, acetate did not accumulate to the same extent. While there was still a rise in concentration over time, at 72 hours, the TMR, ALF, and CEL incubations had 28 mM, 45 mM and 28 mM, respectively (Figure 6.3). The decrease in acetate production is likely due to high H_2 partial pressures (not measured) that make acetate production less thermodynamic favorable.

Acetate's carbon isotope composition became more enriched when methane was
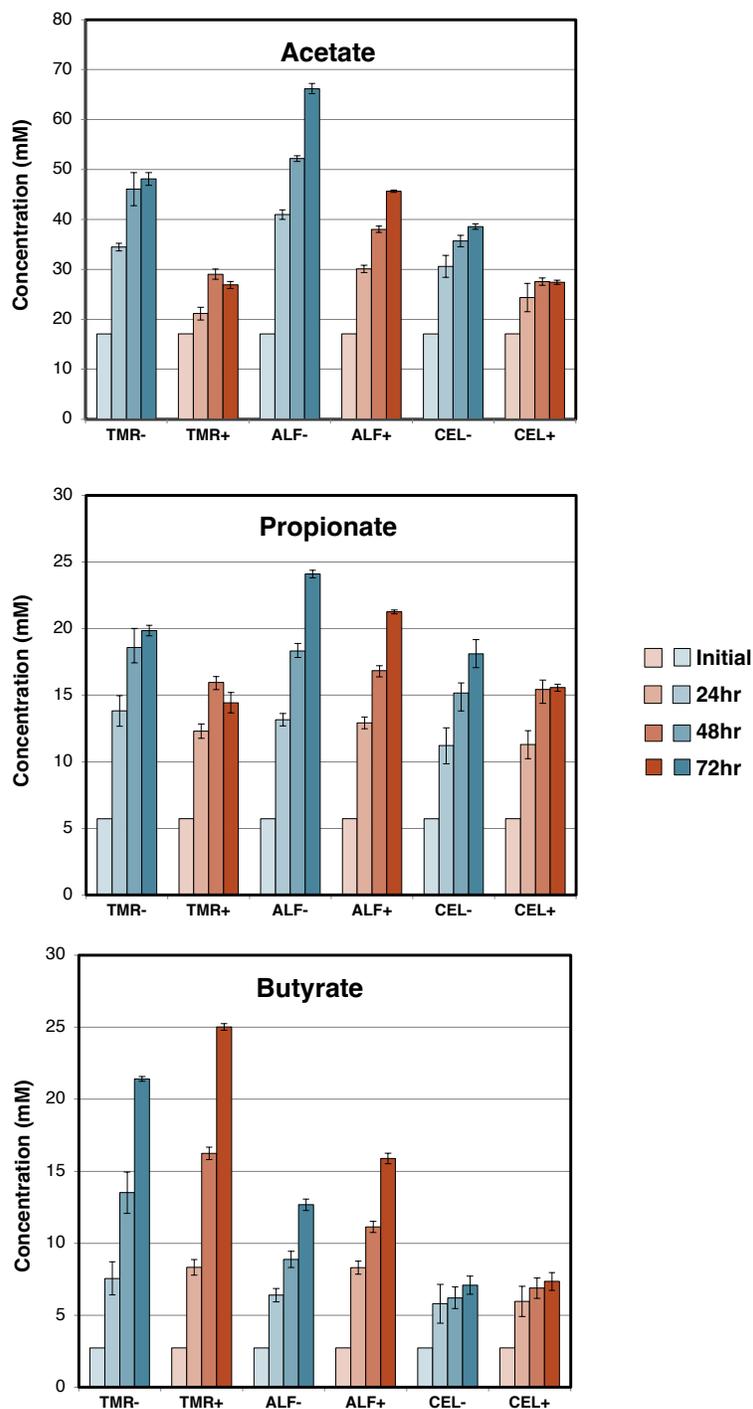
Figure 6.4: VFA concentrations with time over three day rumen incubations. Acetate and propionate production decreased when *A. taxiforms* was added, while butyrate production increased. These trends persisted across all three feeds. The + and - symbols next to the feed condition (e.g. ALF+) indicates conditions with and without *A. taxiforms* , respectively. Abbreviations: ALF, alfalfa; CEL, cellulose; TMR, total mean ration.

inhibited (Figure 6.4, Carbon Number = 2). Initially, acetate was had a $\delta^{13}C$ value of -19‰, about 8‰ enriched relative to the TMR (-27‰). In negative controls , acetate was more ^{13}C-depleted with a $\delta^{13}C$ value of -22 to -26‰. There were no clear correlations with time. However, when *A. taxiforms* was added, acetate in TMR and CEL were closer to the initial value of -19‰, again with no temporal trend. These incubations also had the least acetate accumulation. In the ALF condition, acetate was also enriched but to a lesser extent ($\sim$-21‰). The hydrogen isotope composition of acetate was steady throughout the experiments, shifting between -190 and -210‰, similar to the initial condition.

Propionate

Under all conditions, propionate accumulated in the incubations with small offsets in concentration with and without *A. taxiform* addition. Initial propionate concentrations were $\sim$5 mM and rose to 24 mM, 18 mM and 20 mM with negative controls of TMR, ALF, and CEL, respectively. When *A. taxiforms* was added, total propionate production decreased by 20% relative to negative controls. In the CEL and TMR conditions, the last 24 hours of the experiment had no propionate production, whereas in the ALF condition the rate of propionate increase on day two (24 to 48 hour) was similar to day 3 (48 to 72 hours) (Figure 6.3).

Propionate (Figure 6.4, Carbon Number = 3) had consistently depleted $\delta^{13}C$ composition ($\sim$-26‰) compared to acetate. Its carbon isotope composition did not vary over time, with different feeds, or after methane inhibition. However, the δ^2H composition of propionate changed over time in TMR and ALF conditions when *A. taxiforms* was added to the incubations (Figure 6.4A and B). Propionate's δ^2H value shifted by -30‰ over the course of the experiments (Figure 6.5). However, this trend had a smaller slope when cellulose was the feed source, changing by only -10‰ when *A. taxiforms* was added.

Butyrate

The production rates of butyrate responded differently in our experiments compared to the other VFAs. When *A. taxiforms* was added incubations, net butyrate production in TMR changed from 18.7 mM to 22.3 mM. Similarly, with ALF, net butyrate production increased from 10 mM to 13.1 mM. However, with CEL as the feed source, butyrate production was similar in negative and positive treatments throughout the experiments.

Within 24 hours, the carbon isotope composition of butyrate shifted (Figure 6.4,
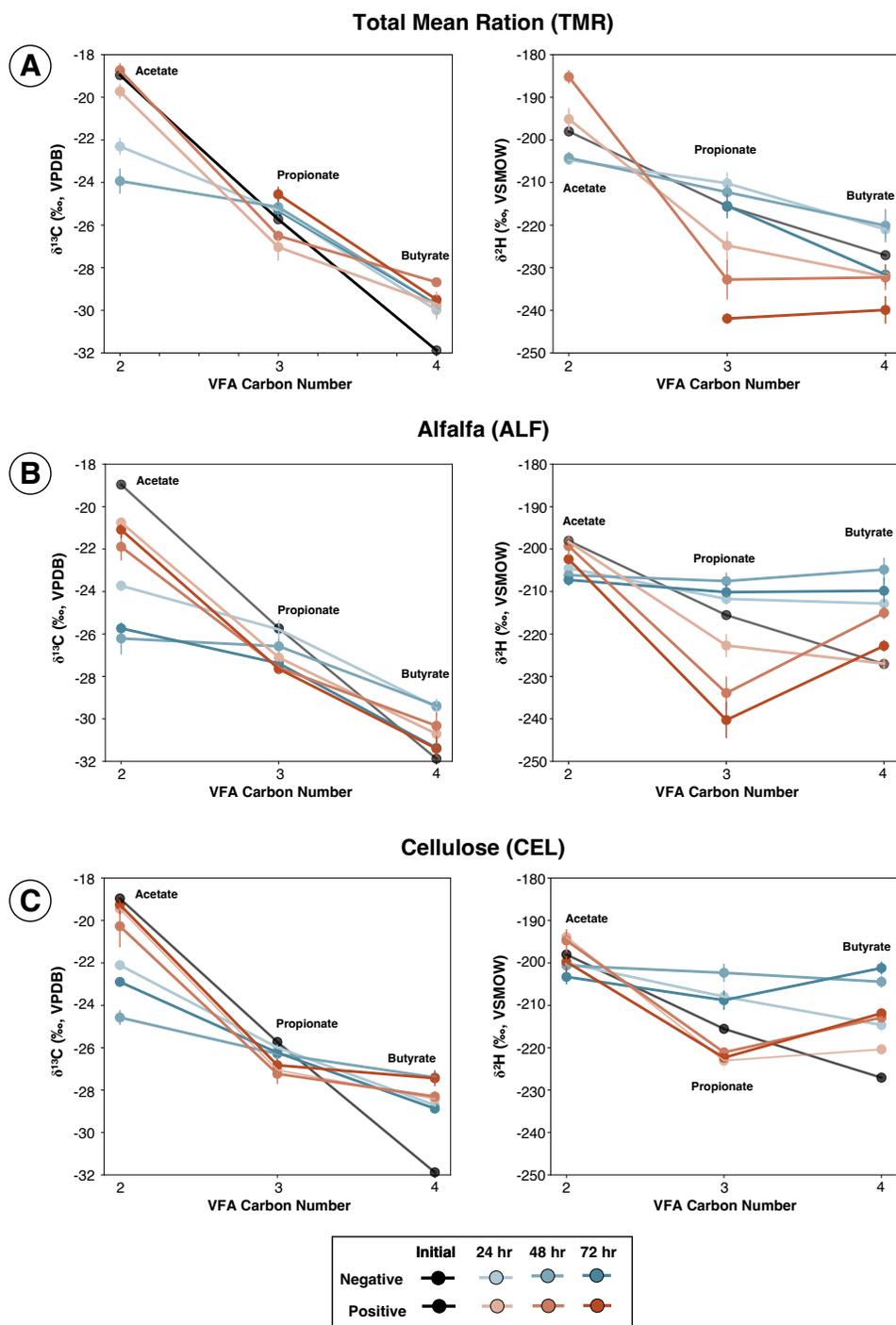
Figure 6.5: The carbon and hydrogen isotope composition of VFAs over three days of rumen incubations plotted against carbon number, with two, three and four representing acetate, propionate and butyrate respectively. These trends are plotted for incubations fed TMR (A), ALF (B) and CEL (C). Black, red and blue lines represent initial conditions, positive treatments with *A. taxiforms*, and negative controls without *A. taxiforms*, respectively. Shading of the color indicates the time with darker hues representing later timepoints.
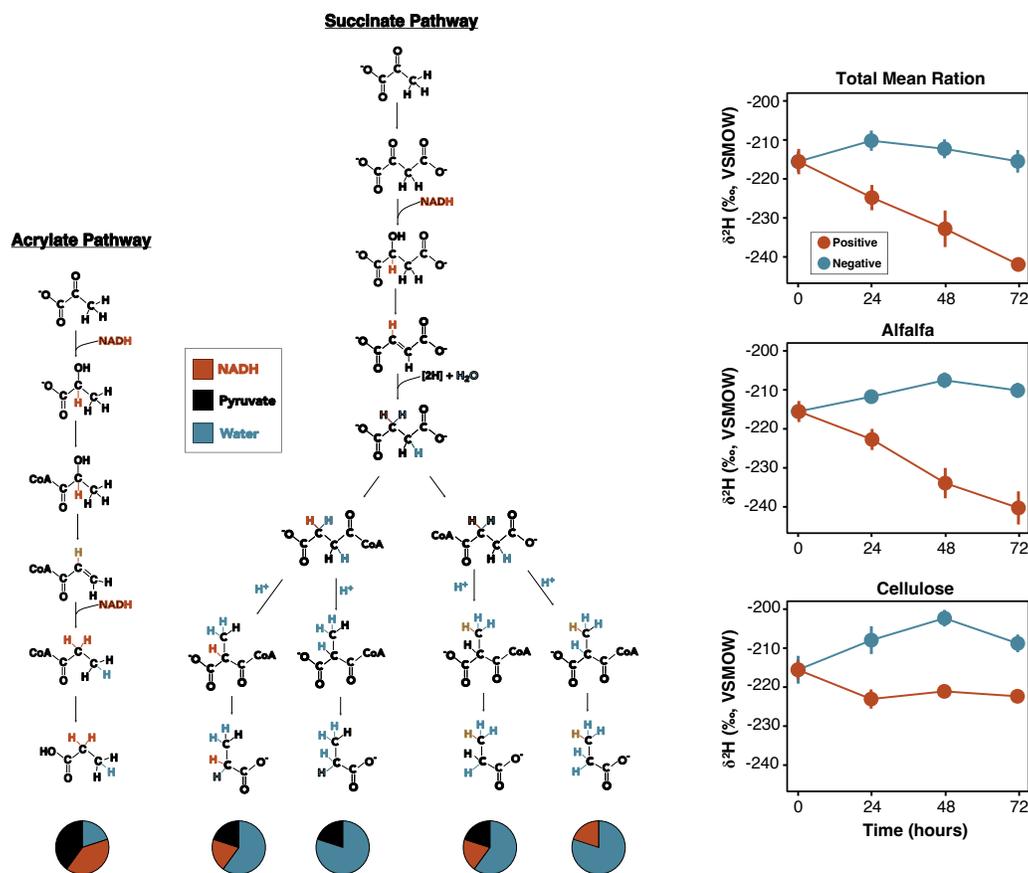
Figure 6.6: On the right, time-varying $\delta^2 H$ values of propionate when incubations were fed different feed types. Blue data points represent negative controls without *A. taxiforms* while red data points represent positive treatments with 2% (w/w) *A. taxiforms* addition. Propionate's $\delta^2 H$ value decreased with time when incubations were fed *A. taxiforms* . This trend did not occur in negative controls. On the left are the two biosynthetic pathways of generating propionate from pyruvate. The three sources of hydrogen atoms — pyruvate, NADH, and water — are color coded black, red, and blue, respectively. Based on the pathway of its formation, propionate inherits its hydrogen atoms from a different ratio of sources, potentially explaining why its $\delta^2 H$ value would shift with the addition of *A. taxiforms* to the feed.

Carbon Number = 4). Initially, butyrate had a $\delta^{13}C$ value of -32‰, which changed to -29‰, -30‰, and -28‰ in the first 24 hours for TMR, ALF, and CEL, respectively (Figure 6.4). This enriched signal held steady for the remainder of the experiment. There were no significant differences between the negative controls and positive treatments. The hydrogen isotope composition of butyrate was consistently ^2H-depleted when *A. taxiforms* was added to the incubation by about 10‰ compared to the negative controls. There was no clear trend with time.

6.5 Discussion

Mitigating methane emissions globally is an important scientific and engineering goal of the next decade. Multiple avenues for decreasing methane production from enteric fermentation in cows have been developed and tested. Given the scale at which these solutions must be implemented, it is important to evaluate their sustainability and efficacy at smaller scales before their widespread use. Regardless of how methanogenesis is inhibited, the rumen microbial community will respond. Just as eliminating an entire species from a forest or lake would profoundly alter the ecosystem's food web, removing methanogenesis from the rumen will have cascading effects on its microbial ecology. Specifically, we were focused on how methane mitigating strategies impact the syntrophy between fermentation and methanogenesis, a metabolic partnership that makes rumen organic degradation so efficient. Their syntrophy is mediated through the exchange of H_2 gas, produced by fermenters and consumed by methanogens. The constant removal of H_2 allows fermenting microbes to continue decomposing plant matter. Since fermentation products from this decomposition are the primary source of energy and carbon for cows, it is crucially important that fermenting microbes continue to process carbon in the absence of methanogens. To adapt, they may upregulate different pathways that are less sensitive to H_2 concentration or that actively consume hydrogen. Probing the mechanisms by which fermenters are responding this new ecological regime will allow us to predict how it will influence cow health and productivity before investing in scaled solutions. To an extent, this is already possible *in vitro* by measuring the rate of VFA production. This has been done for years and has clearly demonstrated that fermentation does indeed respond. However, VFA profiles do not elucidate the metabolic pathways of fermentation nor will they be as useful *in vivo* where the animal host is consuming VFAs and obscuring their production rates.

ESI-Orbitrap isotomics as a new tool for rumen microbiology

Here, we have developed a rapid method to determine the carbon and hydrogen isotope composition of VFA fermentation products. As an analytical tool, ESI-Orbitrap isotope ratio methods were well suited for this study. Since acetate, propionate and butyrate are present in high concentration, minimal preparatory work was required. Samples were filtered, diluted and directly injected onto the instrument through an autosampler. This method uses only 5 μL of rumen fluid. With less than 20 minute analysis times, all six isotopic properties ($\delta^{13}C$ and δ^2H of three VFAs) were measured in the same amount of time (or less) as a liquid or gas chromatography method

that only quantifies their concentrations. In fact, by calibrating the ion currents of each VFA against standard mixtures of VFAs, it may be feasible to quantify both concentration and isotopic composition simultaneously. Thus, ESI-Orbitrap MS represents an exciting avenue for studying not only rumen microbiology but gut fermentation broadly.

Interpreting isotopic signals from VFAs

We next wanted to understand whether VFA isotope compositions could elucidate changes in metabolic pathways caused by the addition of *A. taxiforms* and reduction of methanogenesis. These properties had never been measured before, so we collected VFA concentration, CO_2 and methane concentration and isotope composition data to help interpret our results. More studies are required to fully elucidate the origin of the isotopic signals observed here, but we have developed several impactful and testable hypotheses:

First, our experiments suggest that the $\delta^2 H$ value of propionate could be an indicator of the pathway of its formation. Since propionate is the only VFA that is a net H_2 sink, many investigations have hypothesized that its metabolic formation would change, but there were no clear methods to test that hypothesis. ESI-Orbitrap analyses may fill this gap, serving as a platform to verify the success of microbiome engineering studies attempting to upregulate propionate pathways in the rumen.

Second, we hypothesize that acetate $\delta^{13}C$ composition is an indicator of its production rate. Acetate is quantitatively the most important VFA in the rumen so its production rate is an important factor to characterize when evaluating the sustainability of methane mitigation strategies. In addition, we predict that experiments which successful incorporate homoacetogenesis into the rumen as an alternative H_2 source will have a clear $\delta^{13}C$ signal to distinguish H_2 accumulation (^{13}C-enriched acetate) from H_2 consumption by acetogenesis (^{13}C-depleted acetate).

Below, we detail the evidence behind each of these hypotheses, explore alternative hypotheses that are less likely to explain our data, and discuss the implications of our findings on the rumen carbon flow.

Propionate synthesis pathways shift when *A. taxiforms* is added

Propionate is an important metabolite for redox balancing the rumen microbiome. Fermentation of glucose must balance the production of reducing equivalents like nicotinamide adenine dinucleotide (NADH) from glycolysis with their consumption

during VFA and H_2 synthesis. Propionate, butyrate and acetate syntheses each consume a different number of reducing equivalents. As shown in Equations 6.1-6.4, propionate is the only VFA that uses more reducing equivalents than what is produced during glycolysis (Equation 6.1); it is a net electron sink and consumes 2 moles of H_2 . Butyrate synthesis consumes none of the glycolytic reducing equivalents, generating four moles of H_2 gas. Acetate synthesis, on the other hand, generates two additional reducing equivalents, totalling 6 moles of H_2 per mole of glucose. Cells that produce acetate benefit from additional ATP molecules that are made as a byproduct of its synthesis. To a lesser extent, the same is true for butyrate, while propionate creates the least ATP per mole of glucose. Thus, VFA synthesis is a tradeoff between redox balance and energy generation.

In reality, fermenting cells create multiple VFAs and H_2 simultaneously in a balancing act that maintains redox budget while maximizing ATP yields. This balance changes when methanogens are removed. As H_2 concentration increase, those fermentation pathways that produce H_2 are thought to be less favorable. Indeed, we found that acetate synthesis decreased by almost 50% when *A. taxiforms* was added to the incubations. We found that propionate and butyrate were preferentially synthesized over acetate, likely another adaptation to higher H_2 concentrations. With such low rates of acetate production, ATP yields within fermenting cells must be lower as well. However, the actual ATP budget of rumen fermentation remains unknown, because the ATP yield of propionate synthesis depends on which of two pathways is used, the acrylate or succinate pathway (Figure 6.5). Both pathways start with pyruvate and use two reducing equivalents, but only the succinate pathway generates ATP through a membrane-bound fumarate reductase. Therefore, the actual ATP yield in the rumen is unknown, because we do not know the ratio of these two propionate pathways. This is an important gap to be filled in our understanding of rumen microbial ecology. From an applied perspective, microbiome engineers hope to promote propionate synthesis by incorporating propionic acid bacteria into the rumen. This would shuttle H_2 to more VFA production and lower H_2 concentrations so that fermentation may proceed at a faster pace. These engineering applications would benefit from an understanding of which propionate pathway is best adapted to the new ecology of rumen with methaogenesis inhibited. Differentiating these metabolisms would also serve as a method to verify the success of their microbiome manipulations.

We found that the hydrogen isotope composition of propionate had a clear trend

through time in positive treatments, and this may be a useful indicator of propionate synthesis pathways. To understand the origin of this change, we investigated where propionate inherits its carbon-bound hydrogen. Hydrogen atoms on biomolecules generally have three sources: NADH, the organic substrate, and water. NADH directly transfers hydride ions from its carbon skeleton to that of the biosynthetic product. This has been shown to influence the hydrogen isotope composition of bacterial lipids and amino acids (Wijker et al., 2019; Fogel, Griffin, and Newsome, 2016; Silverman, Wijker, and Sessions, 2024). In the acrylate pathway, lactate dehydrogenase and acrylate-CoA dehydrogenase both use NADH as an electron carrier, which add hydrogen to the propionyl-CoA skeleton (Figure 6.6, red). There is no opportunity for either of these hydrogen atoms to exchange with water once they are carbon-bound (Figure 6.6). However, in the succinate pathway only one of the two reduction steps uses NADH as a carrier. Instead, fumarate reductase transfers hydrogen atoms from water onto succinate. NADH-derived hydrogen make up 40% and 15% of propionate's carbon-bound hydrogen if it is created through the lactate and succinate pathways, respectively. Assuming NADH has a distinct δ^2H value from the pyruvate starting substrate or water, the propionate made from these two pathways would have different δ^2H values. Since the carbon skeleton of pyruvate is preserved in either synthetic route, the $\delta^{13}C$ values of propionate would be similar between the pathways. Thus, we hypothesize that *A. taxiforms* caused the microbiota to shift the proportion of acrylate and succinate pathways used to synthesize propionate, which caused its δ^2H to become more depleted while $\delta^{13}C$ value stayed constant. More experiments with pure cultures of propionic acid bacteria will elucidate which of these pathways is most likely to produce more ^2H-depleted propionate.

We predict that the acrylate pathway will be the best target for these experiments. Hydride transfer is the rate-limiting step of the lactate dehydrogenase reaction that reduces pyruvate with NADH, which causes LDH to have a large KIE (>2.0, k_H/k_D) (Andrés, Moliner, and Safont, 1994). As a result, LDH would transfer an extremely ^2H-depleted hydrogen atom onto propionate. If this is true, rumen microbiota in our experiments partitioned propionate synthesis towards the acrylate pathway during methanogenesis inhibition. In the case of cellulose feed, fermentation rates were less influenced by the addition of *A. taxiforms* than for ALF and TMR, likely because H_2 did not accumulate as quickly. Consistent with this finding, the δ^2H value of propionate separates with time between the CEL incubations with and without *A. taxiforms* but the difference is only 10‰ rather than 30‰ as seen in

TMR and ALF. We note that in the last 24 hours of incubation, propionate in the TMR condition stops accumulating, yet its isotope composition continues to change. This is contradictory to the hypotheses outlined above, unless propionate is being actively consumed and creative. This warrants further exploration.

An alternative hypothesis for the decrease in propionate $\delta^2 H$ value is that the hydrogen isotope composition of NADH became more depleted. This could occur due to shifts in the sources of NADH with increased H_2 partial pressures. However, we would expect that any changes to NADH would also incorporate into acetate and butyrate, but neither of these have a $\delta^2 H$ correlation with time (Figure 6.5). Butyrate's hydrogen isotope compositions broadly shifts to more negative values by 5-10‰ when *A. taxiforms* was added, but no clear trends like those in propionate were observed. If their isotopic depletions were from the same source (NADH), we would expect their isotope compositions to have similar relationships with time and feed condition, which they do not. To test our proposed hypothesis against this alternative, propionic acid bacteria that use either the acrylate or succinate pathways can be grown in pure culture. These experiments would determine the overall $^2 H$-fractionation of the two pathways, which should be different if our proposed hypothesis is correct. If they express a similar fractionation, then the changes in propionate $\delta^2 H$ values found in our experiments is instead a change in the NADH $\delta^2 H$ value.

Acetate $\delta^{13} C$ values reflect shifts towards other VFAs under methanogen inhibition

Acetate is the most abundant VFA in the rumen and a critical source of energy for the animal host. As described above and in Equation 1, acetate synthesis does not consume any reducing equivalents; however, it creates the most ATP per mole of glucose catabolized. When H_2 concentrations accumulate in the absence of methanogenesis, the free energy yield of this reaction rapidly decreases (Equation 5.2). These ideas are consistent with the almost 50% decrease in acetate production with *A. taxiforms* addition. We also observe a shift in the $\delta^{13} C$ composition of acetate from -23‰ to -19‰. These values are 4‰ and 8‰ higher than the carbon isotope composition of the feed, reflecting the isotope fractionation observed in pure cultures in Chapter 4. As discussed in that study, acetate's $^{13} C$-enrichment during fermentation results from several other reactions stemming from acetyl-CoA, the precursor of acetate. Acetyl-CoA is also the precursor of butyrate synthesis, which increases when methanogenesis is inhibited (Figure 6.4). The KIE of acetyl-

CoA acetyltransferase (ACAT) that catalyze the first step of butyrate synthesis has a predicted KIE of 13‰ on both the C2 and C3 sites (see Table 4.1). Shifting acetyl-CoA partitioning from the acetate pathway to the butyrate pathway would cause acetate to become more enriched. The branching ratio of acetyl-CoA would control the magnitude of this enrichment. Thus, we hypothesize that the $\delta^{13}C$ value of acetate compared to that of butyrate and the feed is a metric for the flux of acetate coming from fermentation relative to other VFAs. For example, in the negative control (Blue, Figure 6.5), acetate and butyrate's $\delta^{13}C$ values are offset by 2-4‰, whereas in the *A. taxiforms* conditions (red, Figure 6.5), the offset expands to 6-8‰. This represents a preference for acetate synthesis when fermentation is syntrophically coupled to methanogenesis, but a preference for butyrate when methanogens are inhibited. A lack of acetate synthesis could be harmful to the health and productivity of the animal host. Later, when these strategies are implemented *in vivo*, the rate of acetate production will be a useful metric if animal hosts respond poorly to the seaweed additive.

Acetate $\delta^{13}C$ values as a future proxy for alternative H_2 sinks

Acetogenesis, a metabolism that uses CO_2 and H_2 to generate acetate, is a promising target for alternative metabolic H_2 sink.If acetogens could replace methanogens in the rumen, it would benefit the cattle and dairy industries in three ways simultaneously: It would decrease methane emissions, catalyze fermentation by replacing methanogens with another syntrophic partner, and shuttle more VFA substrate to the animal host. For years, animal scientists and microbiologists have investigated why acetogens are not already a predominant member of the rumen microbial community, given that H_2 concentrations are high enough for them to grow. Some studies have supplemented cow feed with acetogenic spores to encourage their growth. However, at present, acetate produced by fermentation is indistinguishable from acetate produced by reductive acetogenesis. To our knowledge, there are no useful proxies to measure the relative abundance of acetogenic microbes in the rumen or in other environments, including genetic markers. Certain studies have developed primers for genes encoding enzymes in the acetogenesis pathway, but they give no indication of how much acetate is being made through this pathway. In addition, the primers are not specific to acetogenesis, since the gene targets are expressed in other metabolic pathways as well (Gagen et al., 2010). A method to distinguish acetogenesis from other metabolisms would enable us to assess the natural prevalence of acetogenesis across cow rumen samples and to evaluate the success of microbial

supplementation studies.

Carbon isotope compositions of acetate are a useful proxy for determining the relative importance of fermentation and acetogenesis. This has been demonstrated in terrestrial bogs and marine sediments, but has not been applied to the rumen (Gelwicks, Risatti, and J. Hayes, 1989; Blaser, Dreisbach, and Conrad, 2013; Thomas, 2008). In Chapter 1, we determined that Orbitrap analysis of acetate distinguishes acetogenesis from fermentation on two axes: δ^2H and $\delta^{13}C$ values. Thus, the method developed here for rapid quantification of VFA δ^2H and $\delta^{13}C$ values in the rumen would be quite a useful metric for experiments investigating the activity of natural or supplemented acetogenic communities. We expected there to be a shift towards more negative values in the *A. taxiforms* conditions, assuming that acetogens would take advantage of the high H_2 concentrations. However, we see the opposite signal. Given that the background signal of methanogen inhibition is a ^{13}C-enrichment of acetate, detecting a ^{13}C-depletion by acetogenesis will be facile. If acetate's $\delta^{13}C$ value is less than or equal to that of the cow feed, we will know that acetogenesis is contributing. Isotope compositions of acetate can lend information not only about natural shifts in fermentation pathways upon methanogen inhibition but also about the success of engineering tests that hope to replace methanogenesis with a more efficient metabolic sink of H_2 gas.

Possible translation of isotopic tools to *in vivo* studies

Batch incubations of rumen fluid are not fully representative of the *in vivo* conditions within the animal host (E. M. Ungerfeld et al., 2019). Most notably, the host continually absorbs organic acids into their blood to be used for carbon and energy. It also contracts its rumen to move feed around while buffering the pH of its rumen fluid with carbonated saliva (Nagaraja, 2016). Indeed, parallel continuous and batch incubations yield different VFA production rates. During *in vivo* studies, VFA profiles are overprinted by animal host consumption, making them less faithful records of original fermentation pathways and rates. However, we hypothesize that the tools and interpretations built in this study will translate to *in vivo* investigations. The diffusion of organic acids into the host blood stream will possibly impart an isotope effect, which can be quantified. The magnitude of that fractionation is less likely to change between animals or when methanogens are inhibited. Thus, by correcting for the consistent consumption isotopic fractionations, we can quantify the $\delta^{13}C$ and δ^2H values of VFAs produced by fermentation, which in turn will reveal shifts in their metabolic pathways.

6.6 Conclusion

Feed additives and other strategies that reduce methane emissions from enteric fermentation could transform livestock agriculture into more environmentally sustainable industry. However, the implications of these additives on the microbial community that make bovine animals so productive has not been fully elucidated. Current techniques for assessing the rates and mechanisms of microbial fermentation in the rumen cannot distinguish metabolic pathways that create the same fermentation products nor can they assess the *in vivo* rate of VFA generation. We present VFA isotope compositions as a novel tool for filling this gap. ESI-Orbitrap MS methods are well designed for such a task with rapid analysis times, little-to-no sample preparation, and high biological reproducibility. The changes in VFA $\delta^{13}C$ and δ^2H values can yield new information about how fermentation is responding to the inhibition of methanogenesis. Namely, the hydrogen isotope composition of propionate reveals shifts in the propionate synthesis pathways, potentially toward the less energetic acrylate pathway. Meanwhile, the carbon isotope composition of VFAs could indicate the relative abundance of VFAs produced during fermentation of carbohydrates. Finally, our method can, in theory, distinguish acetate produced by acetogenesis from acetate produced by fermentation, presenting an important tool for studies investigating or incorporating reductive acetogenesis in the rumen. Continuous cultures and *in vivo* studies are the next steps to test the hypotheses put forth in our study. We predict that stable isotope measurements will be even more useful in these contexts.

Table 6.2: Composition of artifical saliva buffer

Component	Concentration (mM)
NaCl	28.0
KCl	7.7
$CaCl_2 \cdot 2H_2O$	0.2
$MgCl_2 \cdot 2H_2O$	0.6
NH_4Cl	5.0
$Na_2HPO_4 \cdot 12H_2O$	10.0
$NaH_2PO_4 \cdot H_2O$	10.0
$NaHCO_3$	97.9

6.7 Supplementary Information

The balanced equation for the conversion of glucose to pyruvate with H_2 representing electrons.

$$Glucose + 2H_2O = 2 \times Pyruvate + 2H_2 \tag{6.2}$$
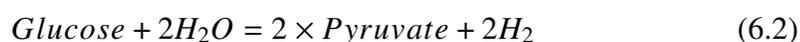
The balanced equation for the conversion of pyruvate to acetate, propionate, and butyrate.

$$2 \times Pyruvate = 2 \times Acetate + 2CO_2 + 2H_2 + 2H^+ \tag{6.3}$$

$$2 \times Pyruvate + 4H_2 = 2 \times Propionate + 2H_2O + 2H^+ \tag{6.4}$$

$$2 \times Pyruvate + 4H_2 = Butyrate + 2CO_2 + H^+ \tag{6.5}$$

To remove biases from sampling and analyses, each Ankom vessel was given a number and the treatment (feed and additive) was randomly assigned to each vessel, while also ensuring four replicates of each treatment. Vessels were sampled during the incubation in order of vessel number. Samples were analyzed (VFA and gas concentration and isotope composition) in order of vessel number to further remove biases.

Table 6.3: Randomization of Ankom vessels and their endpoint pH measured after 72 hours. Positive and negative treatements indicate incubations with and without *A. taxiforms*, respectively. Abbreviations: CEL, cellulose; TMR, total mean ration; ALF, alfalfa

Vessel #	Feed Type	Treatment	Endpoint pH
1	ALF	Positive	5.90
2	ALF	Negative	5.74
3	TMR	Negative	5.26
4	ALF	Negative	5.67
5	CEL	Negative	5.99
6	TMR	Positive	5.95
7	ALF	Negative	5.65
8	ALF	Positive	5.93
9	CEL	Negative	6.06
10	CEL	Negative	5.87
11	CEL	Positive	6.27
12	TMR	Positive	5.53
13	ALF	Negative	5.65
14	CEL	Negative	5.87
15	TMR	Positive	5.43
16	CEL	Positive	6.25
17	CEL	Positive	6.26
18	TMR	Negative	5.23
19	ALF	Positive	6.02
20	TMR	Negative	5.33
21	TMR	Negative	5.10
22	CEL	Positive	6.25
23	TMR	Positive	6.49
24	ALF	Positive	5.94

Table 6.4: Gas production at 24 hour timepoint.

Vessel #	Condition	Total Gas (mL)	CO_2 (mL/g)	CH_4 (mL/g)	CH_4 (‰)	CO_2 (‰)
1	ALF+	210	0.3	36.1	-57.9	-14.8
2	ALF-	168	9.7	14.5	-59.7	-11.9
3	TMR-	200	11.2	17.9	-58.3	-10.3
4	ALF-	180	12.0	17.8	-61.6	-12.2
5	CEL-	122	5.3	9.6	-56.0	-8.6
6	TMR+	160	0.2	18.2	-59.4	-14.2
7	ALF-	240	16.9	31.5	-62.5	-12.6
8	ALF+	172	0.5	25.7	-55.7	-15.1
9	CEL-	126	6.6	12.9	-60.7	-9.8
10	CEL-	150	10.0	14.7	-59.3	-9.2
11	CEL+	105	0.9	8.6	-54.6	-12.4
12	TMR+	240	0.6	43.5	-55.8	-14.8
13	ALF-	210	13.8	22.7	-60.7	-12.3
14	CEL-	131	10.6	15.5	-59.4	-9.2
15	TMR+	192	0.4	52.2	-56.4	-13.8
16	CEL+	128	1.1	17.5	-53.3	-13.2
17	CEL+	127	1.2	13.0	-56.9	-13.1
18	TMR-	196	12.7	24.6	-58.8	-10.9
19	ALF+	158	0.5	22.1	-58.4	-15.2
20	TMR-	182	11.0	20.1	-58.1	-10.8
21	TMR-	205	15.5	37.3	-58.5	-10.6
22	CEL+	131	1.0	21.2	-51.5	-12.1
23	TMR+	205	0.4	34.4	-53.2	-13.2
24	ALF+	148	0.3	21.9	-58.1	-13.9

Table 6.5: Gas production at 48 hour timepoint.

Vessel #	Condition	Total Gas (mL)	CO_2 (mL/g)	CH_4 (mL/g)	CO_2 (‰)	CH_4 (‰)
1	ALF+	161	0.0	29.5	-18.4	—
2	ALF-	182	7.7	25.1	-14.5	-60.8
3	TMR-	175	9.3	29.6	-13.0	-60.6
4	ALF-	186	7.5	27.6	-15.2	-63.3
5	CEL-	96	1.8	12.9	-10.0	-65.5
6	TMR+	187	0.0	36.2	-17.5	—
7	ALF-	190	9.0	37.9	-15.8	-62.4
8	ALF+	151	0.0	28.58	-17.8	—
9	CEL-	92	1.2	8.9	-9.4	-61.5
10	CEL-	91	1.6	12.0	-9.3	-63.1
11	CEL+	85	0.0	9.7	-13.2	—
12	TMR+	210	0.0	49.8	-19.2	—
13	ALF-	185	8.5	34.0	-15.1	-62.7
14	CEL-	110	2.2	14.1	-10.3	-64.2
15	TMR+	182	0.0	69.2	-17.7	—
16	CEL+	66	0.0	7.2	-13.5	—
17	CEL+	76	0.0	8.8	-13.4	—
18	TMR-	180	8.7	33.4	-13.5	-59.0
19	ALF+	168	0.0	38.3	-18.6	—
20	TMR-	209	8.5	36.6	-12.2	-57.5
21	TMR-	169	8.5	38.9	-13.4	-58.9
22	CEL+	80	0.0	11.0	-14.0	—
23	TMR+	182	0.0	48.4	-18.0	—
24	ALF+	157	0.1	34.8	-18.6	—

Table 6.6: Gas production at 72 hour timepoint.

Vessel #	Condition	Total Gas (mL)	CO_2 (mL/g)	CH_4 (mL/g)	CO_2 (‰)	CH_4 (‰)
1	ALF+	111	0.0	11.0	-19.6	—
2	ALF-	134	3.8	16.2	-16.2	—
3	TMR-	64	0.6	4.8	-14.8	—
4	ALF-	132	4.2	20.1	-16.8	-64.9
5	CEL-	41	0.1	1.1	-7.1	—
6	TMR+	136	0.0	22.9	-21.0	—
7	ALF-	125	3.5	15.1	-17.4	-70.4
8	ALF+	131	0.0	16.8	-20.7	—
9	CEL-	51	0.5	2.5	-9.5	-65.3
10	CEL-	49	0.3	2.5	-9.1	-63.4
11	CEL+	41	0.0	1.9	-13.1	—
12	TMR+	136	0.0	20.0	-22.1	—
13	ALF-	139	4.6	18.3	-17.2	-69.8
14	CEL-	52	0.5	3.3	-9.8	-64.2
15	TMR+	136	0.0	19.8	-21.5	—
16	CEL+	45	0.0	2.5	-14.4	—
17	CEL+	47	0.0	2.3	-14.3	—
18	TMR-	103	1.8	12.9	-17.2	-59.0
19	ALF+	109	0.0	17.8	-20.5	—
20	TMR-	105	2.0	14.5	-16.4	-55.5
21	TMR-	102	2.0	13.3	-16.2	-55.4
22	CEL+	42	0.0	2.2	-14.2	—
23	TMR+	128	0.0	19.3	-21.4	—
24	ALF+	116	0.0	22.1	-20.5	—

Table 6.7: VFA $\delta^{13}C$ and δ^2H values at 24 hour timepoint.

		$\delta^{13}C$ (‰, VPDB)			δ^2H (‰, VSMOW)		
Vessel #	Condition	Acetate	Propionate	Butyrate	Acetate	Propionate	Butyrate
1	ALF+	-20.7	-28.1	-32.1	-197	-231	-225
2	ALF-	-23.0	-26.4	-28.9	-203	-214	-212
3	TMR-	-22.2	-25.6	-30.1	-208	-217	-226
4	ALF-	-23.9	-26.9	-30.8	-203	-214	-214
5	CEL-	-21.6	-26.2	-28.6	-195	-214	-217
6	TMR+	-20.7	-28.4	-31.1	-192	-224	-235
7	ALF-	-23.7	-24.9	-29.3	-206	-212	-211
8	ALF+	-21.0	-27.0	-30.6	-198	-219	-231
9	CEL-	-22.4	-25.3	-29.3	-209	-201	-215
10	CEL-	-22.1	-26.5	-27.8	-200	-216	-208
11	CEL+	-19.4	-26.8	-28.3	-197	-224	-218
12	TMR+	-18.7	-27.8	-30.6	-196	-228	-231
13	ALF-	-24.3	-25.0	-28.8	-203	-207	-214
14	CEL-	-22.4	-26.0	-29.1	-195	-201	-219
15	TMR+	-19.7	-26.9	-28.9	-190	-233	-232
16	CEL+	-19.3	-27.7	-28.4	-188	-227	-222
17	CEL+	-20.2	-27.7	-28.3	-194	-226	-224
18	TMR-	-23.4	-25.7	-31.1	-205	-209	-215
19	ALF+	-20.2	-26.5	-29.5	-203	-217	-226
20	TMR-	-22.5	-24.9	-28.7	-201	-213	-220
21	TMR-	-23.2	-25.1	-30.1	-205	-202	-223
22	CEL+	-18.9	-26.1	-28.7	-197	-215	-217
23	TMR+	-19.7	-25.0	-28.2	-203	-215	-230
24	ALF+	-21.2	-26.8	-30.5	-196	-224	-225

Table 6.8: VFA $\delta^{13}C$ and $\delta^2 H$ values at 48 hour timepoint.

Vessel #	Condition	$\delta^{13}C$ (‰, VPDB)			$\delta^2 H$ (‰, VSMOW)		
		Acetate	Propionate	Butyrate	Acetate	Propionate	Butyrate
1	ALF+	-22.0	-26.8	-31.5	-204	-233	-224
2	ALF-	-25.1	-26.1	-28.3	-207	-214	-211
3	TMR-	-22.1	-25.8	-30.0	-206	-212	-227
4	ALF-	-24.4	-26.5	-29.4	-209	-210	-213
5	CEL-	-23.2	-25.3	-27.3	-199	-210	-203
6	TMR+	-19.1	-26.3	-29.1	-185	-233	-225
7	ALF-	-26.9	-27.1	-29.5	-207	-212	-199
8	ALF+	-23.2	-28.5	-29.1	-195	-227	-211
9	CEL-	-24.6	-27.2	-27.8	-200	-200	-208
10	CEL-	-25.1	-27.2	-26.6	-197	-200	-200
11	CEL+	-23.3	-28.0	-28.6	-187	-222	-214
12	TMR+	-19.6	-26.9	-28.5	-181	-221	-228
13	ALF-	-28.3	-27.3	-29.7	-202	-201	-200
14	CEL-	-25.3	-25.5	-28.0	-203	-198	-205
15	TMR+	-17.9	-26.4	-28.6	-186	-229	-239
16	CEL+	-20.6	-26.2	-28.4	-201	-220	-212
17	CEL+	-19.4	-26.4	-28.9	-198	-222	-213
18	TMR-	-23.73	-26.5	-30.0	-202	-212	-226
19	ALF+	-20.7	-27.4	-29.4	-205	-226	-220
20	TMR-	-24.1	-24.5	-30.0	-208	-205	-207
21	TMR-	-25.6	-24.0	-29.3	-201	-219	-221
22	CEL+	-17.9	-28.3	-27.5	-194	-218	-208
23	TMR+	-18.4	-26.4	-28.6	-188	-248	-236
24	ALF+	-23.4	-26.6	-31.3	-200	-215	-214

Table 6.9: VFA $\delta^{13}C$ and δ^2H values at 72 hour timepoint.

Vessel #	Condition	$\delta^{13}C$ (‰, VPDB)			δ^2H (‰, VSMOW)		
		Acetate	Propionate	Butyrate	Acetate	Propionate	Butyrate
1	ALF+	-21.0	-27.8	-30.2	-201	-251	-227
2	ALF-	-25.7	-27.7	-30.0	-206	-207	-209
3	TMR-	-21.6	-24.6	-30.5	-208	-216	-239
4	ALF-	-25.9	-27.4	-31.7	-204	-206	-211
5	CEL-	-23.2	-26.7	-28.7	-205	-211	-203
6	TMR+	-11.3	-23.7	-29.0	-174	-240	-251
7	ALF-	-26.4	-26.7	-31.4	-208	-213	-218
8	ALF+	-20.1	-27.5	-31.5	-202	-227	-221
9	CEL-	-22.8	-26.5	-29.4	-203	-206	-205
10	CEL-	-23.1	-25.8	-28.3	-208	-215	-197
11	CEL+	-20.6	-25.9	-28.4	-200	-224	-210
12	TMR+	-12.5	-24.9	-29.7	-188	-239	-234
13	ALF-	-25.0	-27.8	-32.4	-211	-215	-201
14	CEL-	-22.4	-25.9	-29.1	-198	-203	-200
15	TMR+	-11.7	-25.6	-29.3	-191	-243	-239
16	CEL+	-19.2	-27.0	-27.1	-193	-223	-210
17	CEL+	-19.0	-26.8	-27.9	-201	-224	-214
18	TMR-	-21.4	-26.6	-30.8	-208	-223	-230
19	ALF+	-22.3	-28.1	-33.1	-204	-243	-223
20	TMR-	-20.7	-25.0	-29.1	-208	-207	-229
21	TMR-	-18.4	-25.2	-28.6	-208	-216	-227
22	CEL+	-18.2	-27.6	-26.4	-206	-219	-213
23	TMR+	-13.4	-24.1	-30.0	-186	-245	-236
24	ALF+	-20.9	-27.3	-30.8	-203	-240	-220

BIBLIOGRAPHY

Andrés, J., V. Moliner, and V. S. Safont (1994). "Theoretical kinetic isotope effects for the hydride-transfer step in lactate dehydrogenase". en. In: *J. Chem. Soc., Faraday Trans.* 90.12, pp. 1703–1707. DOI: `10.1039/FT9949001703`.

Beauchemin, K. A. and S. M. McGinn (June 2006). "Methane emissions from beef cattle: Effects of fumaric acid, essential oil, and canola oil1". In: *Journal of Animal Science* 84.6, pp. 1489–1496. DOI: `10.2527/2006.8461489x`. eprint: `https://academic.oup.com/jas/article-pdf/84/6/1489/23651912/1489.pdf`.

Blaser, Martin B., Lisa K. Dreisbach, and Ralf Conrad (Mar. 2013). "Carbon Isotope Fractionation of 11 Acetogenic Strains Grown on H_2 and CO_2". en. In: *Appl Environ Microbiol* 79.6, pp. 1787–1794. DOI: `10.1128/AEM.03203-12`.

Brooke, Charles G. et al. (July 2020). "Methane Reduction Potential of Two Pacific Coast Macroalgae During in vitro Ruminant Fermentation". en. In: *Front. Mar. Sci.* 7, p. 561. DOI: `10.3389/fmars.2020.00561`.

Czerkawski, J.W. (1986). "Energetics of Rumen Fermentation". In: ed. by J.W. Czerkawski, pp. 85–106. DOI: `https://doi.org/10.1016/B978-0-08-025486-9.50015-3`.

Eiler, John et al. (Nov. 2017). "Analysis of molecular isotopic structures at high precision and accuracy by Orbitrap mass spectrometry". en. In: *International Journal of Mass Spectrometry* 422, pp. 126–142. DOI: `10.1016/j.ijms.2017.10.002`.

Etminan, M. et al. (Dec. 2016). "Radiative forcing of carbon dioxide, methane, and nitrous oxide: A significant revision of the methane radiative forcing". en. In: *Geophysical Research Letters* 43.24. DOI: `10.1002/2016GL071930`.

Fogel, Marilyn L., Patrick L. Griffin, and Seth D. Newsome (Aug. 2016). "Hydrogen isotopes in individual amino acids reflect differentiated pools of hydrogen from food and water in *Escherichia coli*". en. In: *Proc. Natl. Acad. Sci. U.S.A.* 113.32. DOI: `10.1073/pnas.1525703113`.

Gagen, Emma J. et al. (Dec. 2010). "Functional Gene Analysis Suggests Different Acetogen Populations in the Bovine Rumen and Tammar Wallaby Forestomach". en. In: *Applied and Environmental Microbiology* 76.23, pp. 7785–7795. ISSN: 0099-2240, 1098-5336. DOI: `10.1128/AEM.01679-10`. URL: `https://journals.asm.org/doi/10.1128/AEM.01679-10` (visited on 05/30/2024).

Gehre, Matthias et al. (May 2015). "On-Line Hydrogen-Isotope Measurements of Organic Samples Using Elemental Chromium: An Extension for High Temperature Elemental-Analyzer Techniques". en. In: *Anal. Chem.* 87.10, pp. 5198–5205. DOI: `10.1021/acs.analchem.5b00085`.

Gelwicks, Jeffrey T., J.Bruno Risatti, and J.M. Hayes (Jan. 1989). "Carbon isotope effects associated with autotrophic acetogenesis". en. In: *Organic Geochemistry* 14.4, pp. 441–446. DOI: 10.1016/0146-6380(89)90009-0.

Glasson, Christopher R.K. et al. (May 2022). "Benefits and risks of including the bromoform containing seaweed Asparagopsis in feed for the reduction of methane production from ruminants". en. In: *Algal Research* 64, p. 102673. DOI: 10.1016/j.algal.2022.102673.

Greening, Chris et al. (Oct. 2019). "Diverse hydrogen production and consumption pathways influence methane production in ruminants". en. In: *The ISME Journal* 13.10, pp. 2617–2632. DOI: 10.1038/s41396-019-0464-2.

Hayes, John M (2001). "Fractionation of the Isotopes of Carbon and Hydrogen in Biosynthetic Processes". en. In.

Hobson, P. N. and C. S. Stewart, eds. (1997). *The Rumen Microbial Ecosystem*. en. Dordrecht: Springer Netherlands. DOI: 10.1007/978-94-009-1453-7.

Johnson, K. A. and D. E. Johnson (Aug. 1995). "Methane emissions from cattle". In: *Journal of Animal Science* 73.8, pp. 2483–2492. DOI: 10.2527/1995.7382483x. eprint: https://academic.oup.com/jas/article-pdf/73/8/2483/22737690/2483.pdf.

Kung L., Jr. et al. (Jan. 2003). "Effects of 9,10 anthraquinone on ruminal fermentation, total-tract digestion, and blood metabolite concentrations in sheep1". In: *Journal of Animal Science* 81.1, pp. 323–328. DOI: 10.2527/2003.811323x. eprint: https://academic.oup.com/jas/article-pdf/81/1/323/23748610/323.pdf.

Machado, Lorenna et al. (Jan. 2014). "Effects of Marine and Freshwater Macroalgae on In Vitro Total Gas and Methane Production". en. In: *PLoS ONE* 9.1. Ed. by Douglas Andrew Campbell, e85289. DOI: 10.1371/journal.pone.0085289.

Maia, Margarida R. G. et al. (Aug. 2016). "The Potential Role of Seaweeds in the Natural Manipulation of Rumen Fermentation and Methane Production". en. In: *Sci Rep* 6.1, p. 32321. DOI: 10.1038/srep32321.

Makarov, Alexander and Eduard Denisov (Aug. 2009). "Dynamics of ions of intact proteins in the Orbitrap mass analyzer". en. In: *J. Am. Soc. Mass Spectrom.* 20.8, pp. 1486–1495. DOI: 10.1016/j.jasms.2009.03.024.

Mueller, Elliott P. et al. (Jan. 2022). "Simultaneous, High-Precision Measurements of δ ^{2} H and δ ^{13} C in Nanomole Quantities of Acetate Using Electrospray Ionization-Quadrupole-Orbitrap Mass Spectrometry". en. In: *Anal. Chem.* 94.2, pp. 1092–1100. DOI: 10.1021/acs.analchem.1c04141.

Nagaraja, T. G. (2016). "Microbiology of the Rumen". In: ed. by Danilo Domingues Millen, Mario De Beni Arrigoni, and Rodrigo Dias Lauritano Pacheco, pp. 39–61. DOI: 10.1007/978-3-319-30533-2_2.

National Academies of Sciences, Engineering and Medicine (July 2018). *Improving Characterization of Anthropogenic Methane Emissions in the United States*. en. Pages: 24987. Washington, D.C.: National Academies Press. DOI: 10.17226/24987.

Patra, Amlan et al. (Dec. 2017). "Rumen methanogens and mitigation of methane emission by anti-methanogenic compounds and substances". en. In: *J Animal Sci Biotechnol* 8.1, p. 13. DOI: 10.1186/s40104-017-0145-9.

Penning, Holger and Ralf Conrad (May 2006). "Carbon isotope effects associated with mixed-acid fermentation of saccharides by Clostridium papyrosolvens". en. In: *Geochimica et Cosmochimica Acta* 70.9, pp. 2283–2297. DOI: 10.1016/j.gca.2006.01.017.

Romero, Pedro et al. (Nov. 2023). "Rumen microbial degradation of bromoform from red seaweed (Asparagopsis taxiformis) and the impact on rumen fermentation and methanogenic archaea". en. In: *J Animal Sci Biotechnol* 14.1, p. 133. DOI: 10.1186/s40104-023-00935-z.

Roque, Breanna Michell et al. (Dec. 2019). "Effect of the macroalgae Asparagopsis taxiformis on methane production and rumen microbiome assemblage". en. In: *anim microbiome* 1.1, p. 3. DOI: 10.1186/s42523-019-0004-4.

Silverman, Shaelyn N., Reto S. Wijker, and Alex L. Sessions (Apr. 2024). "Biosynthetic and catabolic pathways control amino acid δ2H values in aerobic heterotrophs". en. In: *Front. Microbiol.* 15, p. 1338486. DOI: 10.3389/fmicb.2024.1338486.

Smith, Pete et al. (2011). "11 Agriculture, Forestry and Other Land Use (AFOLU)". en. In.

Stolper, D.A. et al. (Feb. 2014). "Combined 13C–D and D–D clumping in methane: Methods and preliminary results". en. In: *Geochimica et Cosmochimica Acta* 126, pp. 169–191. ISSN: 00167037. DOI: 10.1016/j.gca.2013.10.045. URL: https://linkinghub.elsevier.com/retrieve/pii/S0016703713006170 (visited on 04/16/2024).

Thomas, Randal B (2008). "Intramolecular isotopic variation in acetate in sediments and wetland soils." en. In.

Ungerfeld, Emilio (Apr. 2020). "Metabolic Hydrogen Flows in Rumen Fermentation: Principles and Possibilities of Interventions". en. In: *Front. Microbiol.* 11, p. 589. DOI: 10.3389/fmicb.2020.00589.

Ungerfeld, Emilio M. (Feb. 2015). "Shifts in metabolic hydrogen sinks in the methanogenesis-inhibited ruminal fermentation: a meta-analysis". en. In: *Front. Microbiol.* 6. DOI: 10.3389/fmicb.2015.00037.

Ungerfeld, Emilio M. et al. (Apr. 2019). "Inhibiting Methanogenesis in Rumen Batch Cultures Did Not Increase the Recovery of Metabolic Hydrogen in Microbial Amino Acids". In: *Microorganisms* 7.5, p. 115. DOI: 10.3390/microorganisms7050115.

White, D., J.T. Drummond, and C. Fuqua (2012). *The Physiology and Biochemistry of Prokaryotes*. Oxford University Press.

Wijker, Reto S. et al. (June 2019). "^{2}H/^{1}H variation in microbial lipids is controlled by NADPH metabolism". en. In: *Proc. Natl. Acad. Sci. U.S.A.* 116.25, pp. 12173–12182. DOI: `10.1073/pnas.1818372116`.

*

FROM STANDARDS TO SAMPLES: BEST PRACTICES FOR ANALYZING ORGANIC ACID ISOTOPE COMPOSITIONS FROM THE ENVIRONMENT USING AN ESI-ORBITRAP.

Elliott P. Mueller[1*], John Eiler[1], Alex L. Sessions [1]

[1]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA

Mueller, Elliott et al. (2024). "Isotope ratio analysis of acetate using Orbitrap Exploris Isotope Solutions". In: *Thermo Fisher Technical Note* Isotope ratio MS.
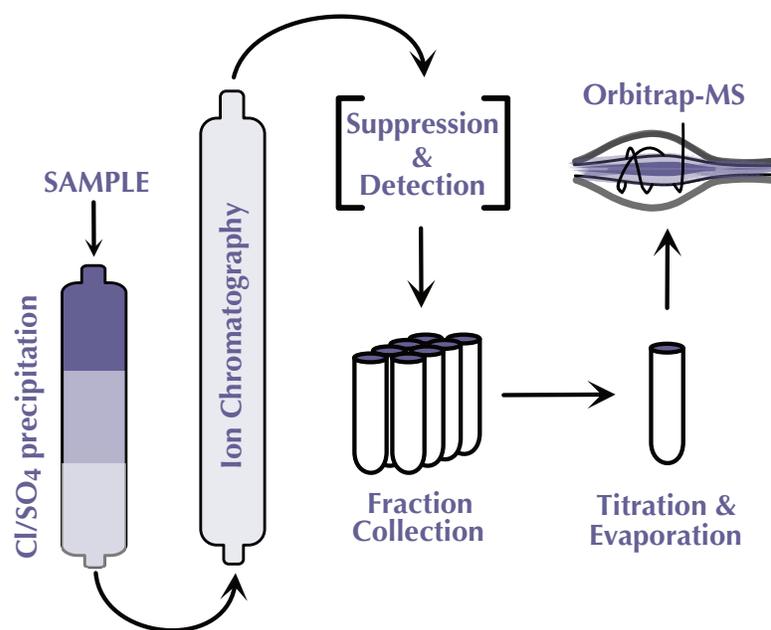
Graphical Abstract of Chapter 7: Workflow of organic acid purification from an environmental sample.

7.1 Abstract

In the past decade, bioanalytical mass spectrometry (MS) has advanced to measure stable isotope ratios in organic molecules at their natural abundances. These developments have substantially broadened the scope of molecular and isotopic targets available to address questions in the natural sciences. However, research has focused on optimizing measurements of pure organic molecules, known as 'standards'. To transition our methods from analyzing pure standards to environmental samples, we must account for how the matrix in which target analytes are dissolved can affect instrument performance and alter measured isotope ratios. In this study, we utilized organic acids as a case study to investigate these 'matrix effects'. Our findings reveal that major anions such as chloride can greatly suppress analyte signals during electrospray ionization. In response, we developed a specialized workflow for isolating organic acids from environmental samples. This method minimizes ion suppression and enables accurate $\delta^{13}C$ and $\delta^2 H$ measurements, even in saline brines. A second major issue was limiting sample consumption rates during analysis. We outline how autosampler "in flow" injections can minimize sample usage while maximizing throughput. Samples can also be conserved by adopting a quadrupole "shifting" technique that measures the $\delta^{13}C$ and $\delta^2 H$ values of multiple molecular targets within a single acquisition. However, even after purification and quadrupole filtering, contaminant ions can appear in mass spectra. In our samples, contaminants altered carbon isotope ratios of acetate by >20‰. We developed strategies for eliminating contaminants in sample preparation and during ionization. Together, these studies provide a framework for observing, mitigating, and testing matrix effects during ESI-Orbitrap measurements, one that should be continually revisited as we transition from pure standards to natural samples.

7.2 Introduction

The advent of bioanalytical mass spectrometry for stable isotope applications is revolutionizing the field of stable isotope biogeochemistry. These techniques are not only unlocking the high-dimensional properties of molecular isotopic structures but also expanding the diversity of molecular targets amenable to such analyses. Previously, biomolecules like amino acids, organic acids, lipids, sugars, and secondary metabolites had to be chemically converted to small gases like CO_2 and H_2 before measuring their molecular-average $\delta^{13}C$ and $\delta^2 H$ values. This practice eliminated the dense isotopic information stored within molecules themselves such as site-specific isotopic enrichments and nonstochastic distributions of multiply-

substituted isotopologues (J. Eiler et al., 2017; Zeichner, Aponte, et al., 2023; Csernica, Sessions, and J. M. Eiler, 2023). The key feature of bioanalytical MS that differentiates it from conventional isotope ratio MS (IRMS) is its analysis of intact molecular and fragment ions. Analytes are infused directly into the mass spectrometer without chemical conversion, charged by electrospray ionization, and then analyzed by high-resolution mass spectrometry (e.g. time-of-flight, Orbitrap, ion cyclotron).

This advancement will substantially expand the scope of stable isotope chemistry. Chapter 1 of this thesis represents one step into such an expanse, demonstrating that hydrogen isotopes, the rarest of the commonly used stable isotopes, can be measured precisely and accurately via ESI-Orbitrap MS. Similar steps forward have been made by my colleagues at Caltech on molecules from amino acids like methionine and alanine to sugars like gluconate to inorganic anions like sulfate (Csernica, Sessions, and J. M. Eiler, 2023; Neubauer, Sweredoski, et al., 2018; Neubauer, Crémière, et al., 2020; Hilkert et al., 2021; Weiss et al., 2023). This represents only a small subset of methods being developed globally. However, these pioneering studies have overwhelmingly focused on pushing the limits of Orbitrap MS in terms of sensitivity, ion counting efficiency, internal precision, and long-term reproducibility (Csernica, Bhattacharjee, and J. Eiler, 2023; Kew et al., 2023; Mueller, Sessions, et al., 2022; Weiss et al., 2023; Wang et al., 2024). At present, applications of these techniques to natural systems are less common. Chapters 5 and 6 of this thesis demonstrates two of the first use cases of Orbitrap MS for environmental measurements. In preparing and executing these studies, it became clear to me that a large hurdle for this technology and the scientists shepherding it into the broader community will be the translation of our methods — optimized for pure standards — to more complex natural samples. The following results represent a series of lessons I learned while attempting to surmount this issue.

Such challenges have long been recognized by those using gas chromatography coupled to Orbitrap MS (GC-Orbitrap) for isotope ratio analyses. However, the GC-Orbitrap platform benefits from on-line separation of analytes via GC before analysis, whether in direct elution or long term reservoir injections (Hofmann et al., 2020; Wilkes and Pearson, 2019; Zeichner, Wilkes, et al., 2022; Zeichner, Aponte, et al., 2023; Chimiak et al., 2021). Furthermore, both gas chromatography and electron impact ionization have decades of research to contextualize GC-Orbitrap applications. In fact, these technologies have been coupled for over forty years to

make compound-specific isotopic analyses via GC-IRMS. ESI-Orbitrap may one day be coupled directly to liquid chromatography (LC) to gain a similar advantage of on-line separation. At present though, most measurements are made by direct infusion of a sample into the Orbitrap. This means that analytes must be completely purified from their environmental matrices before analysis or be injected along with their matrices. Unlike our counterparts in GC-Orbitrap research, our studies are the first to use liquid-source mass spectrometers (e.g. ESI, APCI) for isotope ratio analysis and thus, many questions remain unanswered regarding the robustness of these instruments to chemical interferences (Hilkert et al., 2021; Neubauer, Crémière, et al., 2020). For example, how will the co-existence of target analytes with matrix ions like chloride influence the efficacy of the ESI-Orbitrap's ion source, quadrupole mass filter, ion trap, and mass analyzer? Will these changes affect the isotope ratio of the analyte that we are trying to measure? Matrix effects are an ever-present issue in analytical chemistry. The studies presented here are not comprehensive and cannot be applied to all ESI analytes. Instead, they outline a path forward for testing, mitigating and correcting such effects.

7.3 Methods

For precipitation of chloride and sulfate, samples were run through a Dionex Ag/Ba/H cartridge at 0.5 mL/min to remove chloride after the cartridge had been washed with 300 mL purified water (Milli-Q) at 2 mL/min. The first 0.5 mL of eluent from the cartridge was discarded as it represented the dead volume. The remaining sample was collected until almost all the resin was used, carefully avoiding over-filling the cartridge, which would cause chloride to leak through.

The compound-specific isotope composition ($\delta^2 H$ and $\delta^{13}C$) of acetate, propionate and butyrate were measured on a heated electrospray ionization (HESI) Orbitrap mass spectrometer (MS) (Thermo Scientific) attached to a Vanquish HPLC (Thermo Scientific) with a split-sampler and 100 μL sample loop. No LC column was used for these analyses. Direct infusion and dual inlet measurements used a Fusion 101 (Chemyx, Inc.) syringe pump equipped with 500 μL syringe (Hamilton Robotics). The pump infused sample into the Orbitrap at 5 μL/min. For autosampler injections, 100 μL of sample was pulled through the sample loop and then injected into the 5 μL/min flow of LC-MS grade methanol (Fisher chemical, Optima), which carried the sample directly into the Orbitrap MS. The ESI ionization parameters are the same as in Table 6.1. The raw mass spectra were converted to ion counts using the Makarov equation. The signal-to-noise of ion intensities (S/N$_P$) from mass spectra

were converted to counts (N) by applying an empirical factor (C_N = 4.4) derived by Makarov and Denisov (2009) and the charge of the ion (z). Isotope ratios were first analyzed on the "IsoXL" online software and then processed on a bespoke Python pipeline for data processing as in Chapters 5 and 6.

$$N = \left(\frac{S}{N_P} \right) \left(\frac{C_N}{z} \right) \sqrt{\frac{R_N}{R}} \tag{7.1}$$

7.4 Results and Discussion

My chosen molecular targets for Orbitrap isotope ratio analysis are organic acids. For the majority of my work, I specifically measured the C_2 acid, acetate. Acetate is particularly difficult to ionize compared to other small organics like amino acids and sugar acids. At 50 μM, acetate in pure solution only has an ion current of $\sim 10^8$ NL score (a dimensionless unit that scales with the ion current), whereas other molecules achieve $>10^9$ NL scores. To gain even this level of ionization, the auxiliary gas of the ESI source has to be heated to 100°C. Notably, heated electrospray (HESI) is not available on all ESI-Orbitraps. Given its lackluster ionization efficiency, it does not do well in competition with other ions in solution. In Chapter 2, I found that a 20-to-1 ratio of chloride with acetate nearly eliminates its ionization entirely (Figure 2.6). Unfortunately, in nature, acetate is found at even higher ratios with chloride (Figure 7.1). Thus, even to observe acetate, it had to be isolated from these interfering ions.
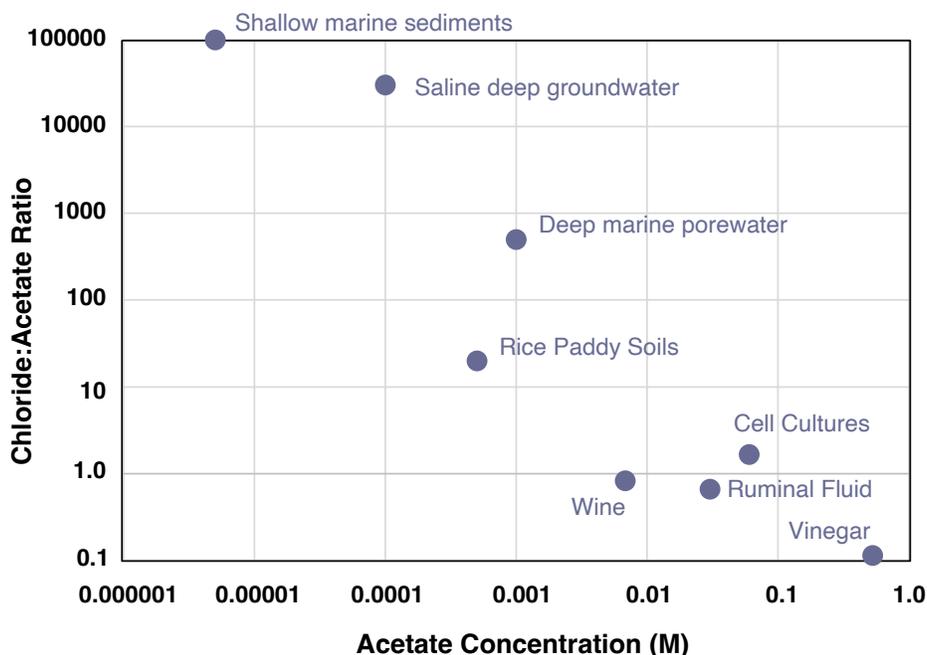
Figure 7.1: Acetate concentrations in natural samples have a wide range from 100mM to 1μM. The ratio of chloride to acetate concentrations also varies but it rarely less than one. To measure acetate on electrospray bioanalytical mass spectrometers, it must be isolated from these interfering ions, which suppress ionization — like chloride. Data from: (Sherwood Lollar et al., 2021; Thomas, 2008; Ungerfeld, 2015; Mueller, Wu, and Sessions, 2022; Beulig et al., 2018; Heuer et al., 2020; Liesack, Schnell, and Revsbech, 2000)

Isolating Acetate from Matrices

Acetate (and other organic acids) can successfully be isolated with a combination of preparatory precipitation chemistry and ion chromatography. This procedure likely applies to other organic or inorganic anions. High performance ion chromatography (HPIC) with an anion exchange stationary phase and basic mobile phase (e.g. potassium hydroxide) can easily separate organic acids (green, Figure 7.2) from inorganic anions (purple, Figure 7.2). Cations are not retained on the column and elute well before the organic acids. The advantage of this approach is that the mobile phase is only water and potassium hydroxide (KOH). The potassium cations are replaced with hydronium ions as the eluent passes through the electrolysis suppressor. The analytes are then carried by pure, slightly acidic water. The concentration is non-destructively quantified by conductivity detection and the eluent can be diverted to a fraction collector (in this case my hand was said collector). As a result, analytes are collected in pure water as opposed to organic or ionic eluents used commonly in

LC separation. This is ideal for Orbitrap analyses, which may require pure samples.
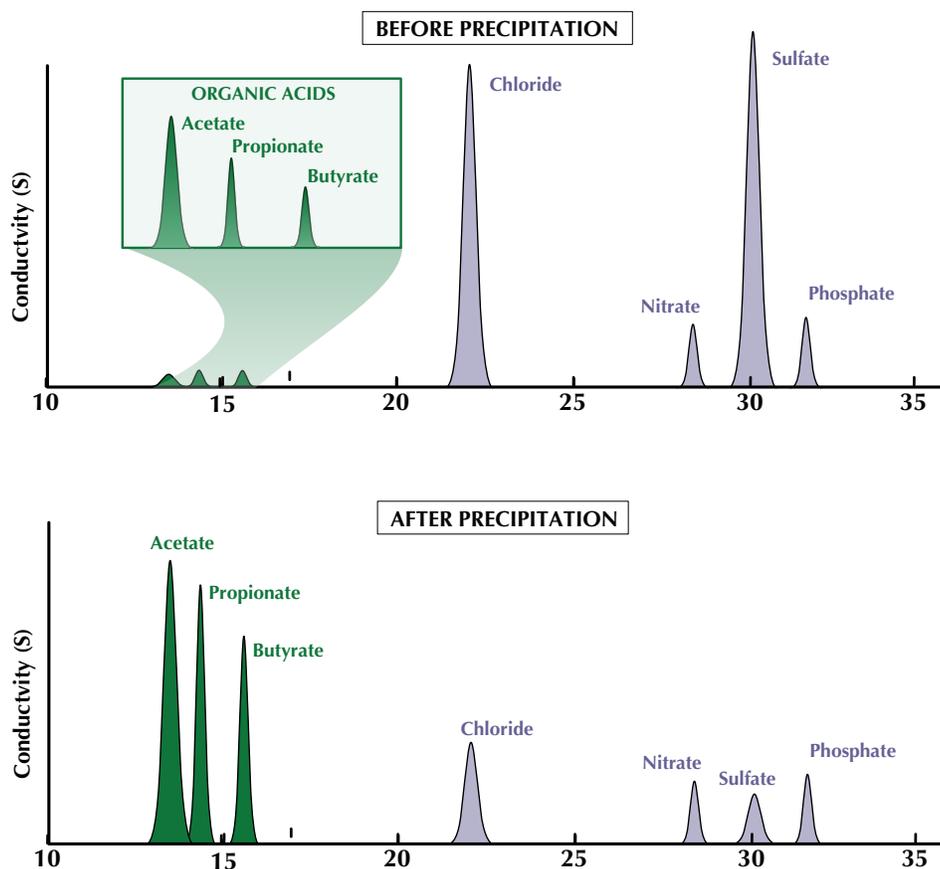
Figure 7.2: Schematic of an ion chromatogram when low volumes of diluted sample are injected onto the HPIC column. Organic acids are easily resolved from inorganic anions and oxyanions. This highlights the advantages of HPIC in preparing compounds for ESI-Orbitrap analysis.

The major drawback of ion chromatography is its column capacity. Low volumes of diluted sample (10 μL) can be injected with good peak resolution. However, to collect enough sample for Orbitrap analysis (~100 nanomoles), large volumes of undiluted sample must be loaded onto the HPIC. In practice, this overloads the ion exchange capacity of the column with the major anion of the matrix (usually chloride) and the organic acids cannot retain to the stationary phase and are lost. To circumvent this issue, the major anions are removed before HPIC using silver and barium resins within flow-through cartridges that precipitate silver chloride and barium sulfate. These single-use cartridges effectively remove high concentrations of chloride and sulfate without losing much sample or diluting it. However, they had

significant blank contributions. With pure water flowing through them, no acetate could be observed in the eluent. When the resin was partially used by a pure solution of sodium chloride, acetate started to leak out of the cartridge, presumably bound to the resin and released during its conversion to silver chloride. By flushing the resin with excessive volumes of purified water (>300 mL), this contamination was removed. With this issue resolved, the cartridges were almost 100% effective in removing chloride and sulfate, but the efficacy inversely correlated with the flow rate. When samples were flowed at <0.5 mL/min, chloride removal was nearly complete, even in brines with 3.5 M chloride. If flow rates exceed 1 mL/min, chloride leaks through, particularly when the cartridge is overused and the resin is exhausted. Thus, a common rule was to only use 90% or less of the cartridge's resin (observed as a color change) in order to avoid leaking chloride to the sample.

With the major ions removed, the samples can be injected directly onto HPIC columns without dilution. At such high injection volumes, the organic acids elute as a co-eluting multi-modal chromatograph peak that is collected in a single vial. However, it is sufficiently separated from chloride and sulfate. The resulting solution is about pH 4 when it elutes, since it is degassed, pure water. At that pH, acetate becomes its protonated form acetic acid, which is highly volatile. To avoid further sample loss, it is titrated with 5 μL of 0.1M NaOH to pH 7-8. If the solution becomes too basic, CO_2 from the atmosphere will dissolve into the sample and create issues during mass analysis (see Contaminant section). The sample is then dried down at room temperature over a flow of pure nitrogen gas and then dissolved into methanol to an approximate concentration of 50 μM.

We tested this workflow (Chapter 7 Graphical Abstract) on a synthetic brine solution spiked with acetate. This solution had a chloride-to-acetate ratio of 3500-to-1. The total yield of the procedure was only 30-40% over several replicates. Acetate was likely lost during the precipitation chemistry and the final dry down steps. Yield could significantly be improved if acetate could be ionized in water or water-methanol mixtures which would avoid sample dry down. However, at present, that is not possible.

The low yield did not fractionate the isotopic composition of acetate. The acetate spiked into these solutions had a known isotope composition, which was accurately measured after acetate was extracted from the brines and analyzed (Figure 7.3). Furthermore, we analyzed eight natural samples of deep continental saline groundwaters, each with a salinity comparable to the synthetic mixture. All of these
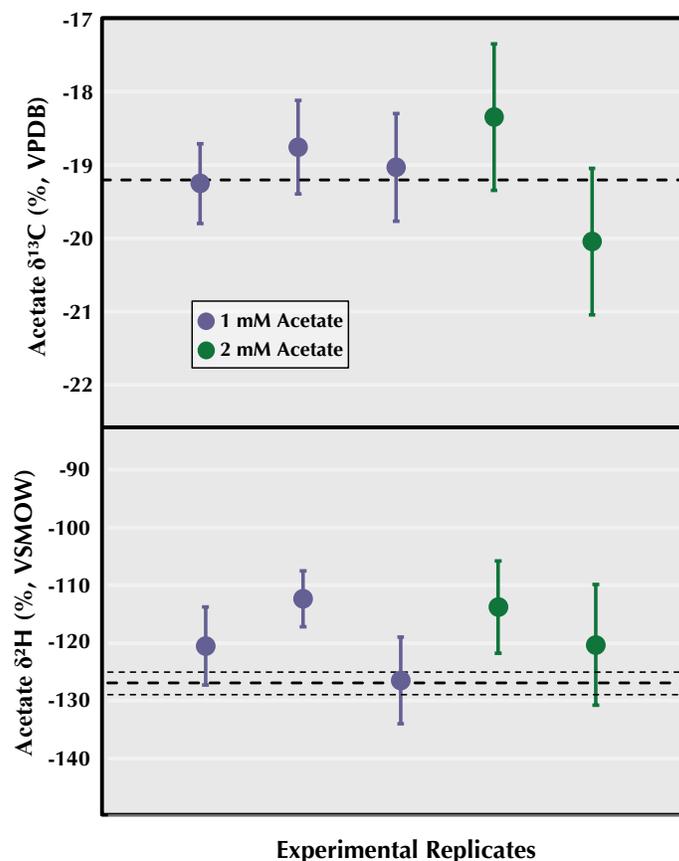
Figure 7.3: Analysis of acetate with a known isotope composition ($\delta^{13}C$ = -19.2‰, δ^2H = -127‰) that had been extracted from a synthetic brine solution (full description of brine solution can be found in Table 5.1). The extraction procedure does not fractionate acetate's isotope composition despite <50% yields.

samples had previously been measured by LC-IRMS for the $\delta^{13}C$ value of acetate. The Orbitrap and LC-IRMS techniques agreed nicely (Figure 7.4). Thus, we feel confident that even in highly saline brines, the sample workflow can measure the $\delta^{13}C$ and δ^2H of acetate accurately.

Sample Introduction Systems

When working with samples, mitigating sample consumption is a priority. The amount of sample required per analysis is a function of flow rate, total analysis time, and the fraction of analysis time during which sample is being consumed. The latter two parameters are determined by the sample introduction system. During my thesis, I developed three separate methods for sample introduction. Each of these can attain the requisite precision and accuracy, but they differ in their throughput and
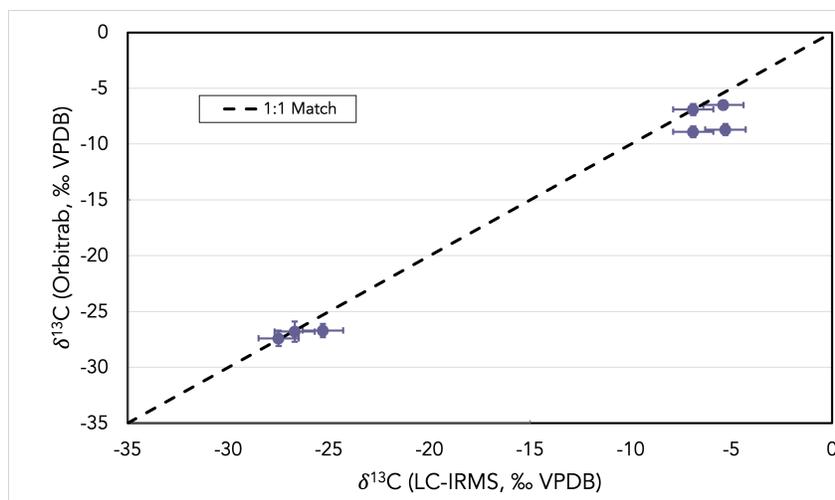
Figure 7.4: A further test of the fidelity of acetate extraction procedures. Comparing LC-IRMS analyses of brine solutions, which has on-line separation methods to purify acetate, to the off-line purification and ESI-Orbitrap technique used here, we find good agreement between the techniques. Hydrogen isotope accuracy can only be measured by spiking standards into synthetic solutions, as environmental measurements of acetate's δ^2H value is possible at present by other techniques.

sample consumption rates. In Tables 7.1 and 7.2, I outline a comparison of these techniques. Direct infusion methods are the fastest, as almost all of the time goes to sample analysis. Flushing between samples is done manually with clean solvents, which makes it fast and uses minimal sample. However, it requires the user to be present for all analyses. The dual inlet method is slightly more automated, though it is the one method that uses sample solution for the entirety of the method, even when standards are being run. It also takes 1-5 minutes in practice to fully switch between sample and standard, further increasing sample consumption rates. Only about 30% of the sample used in this method is used in the final data analysis.

The autosampler or "in-line" introduction system combines automation and minimal sample consumption. For a 10 minute acquisition, 50-55 μL of sample is pulled into a sample loop and infused into the MS. Unlike direct infusion or dual inlet measurements, almost no sample is wasted and nearly all of it is used in the final data analysis. After the acquisition, pure solvent from the reservoirs in the LC are used to purge the system.This takes about 10 minutes but decreases the background acetate signal significantly. By increasing the flow rate during the flushing stage, even lower background signals can be achieved. With this method only half of the total time (140 minutes) is used to analyze sample solutions, but this decreased efficiency is easily compensated for by the automation which allows the user to

analyze samples continuously for days if not weeks. All the data in Figure 7.6 was collected using the autosampler in a continuous run that lasted 1.5 weeks with minimal interruptions for routine maintenance. In comparing these sample introduction techniques, autosampler injections are the clear choice for organic acid analyses, largely because they increase throughput and conserve sample as much as possible while removing elements of human error.

Ion Suppression

The preceding discussion focused on acetate, but many other organic acids exist in nature and serve as important intermediates in the carbon cycle. The isolation workflow is agnostic to the exact organic acid target, as it collects all of them in a single fraction. While this means more organic acids can be evaluated on the ESI-Orbitrap, these mixtures are their own kind of matrices, albeit simpler than the original sample's matrix. In our analyses of brine solutions from the deep continental subsurface, we found that acetate's ion current was lower than we would expect given the concentration of acetate in the final, purified solution (50 μM - 100 μM). Relative to standards, the ionization yield of acetate in samples was lower. This may have been due to ion suppression by formate, another organic acid that co-elutes with acetate off the HPIC column. We tested this idea on combinations of sodium acetate (C_2), sodium propionate (C_3) and sodium butyrate (C_4) in pure methanol. Holding acetate's concentration constant at 50 μM, the concentrations of the other two organic acids were adjusted to alter the relative abundance of acetate in solution from 20% to 100% (Figure 7.5). The ion current of acetate (measured with the quadrupole mass filter set to 57 — 62 m/z) had a linear relationship with acetate's relative abundance. The other organic acids were suppressing acetate's ionization, outcompeting it. The ion current dropped by 3-fold between solutions. However, the $\delta^{13}C$ and δ^2H values of acetate measured relative to a standard of pure 50 μM acetate were consistently within two standard deviation of the true value. Thus, ion suppression did not fractionate acetate's isotopic composition.

Mixing of sample with background ('blank') acetate can cause isotopic fractionations. At the lowest relative abundance, the ion current was only 2 to 5-times higher than the blank (1-3 $\times 10^7$, Figure 7.5). As shown in Figure 2.5, the contribution of blank can alter the measured δ^2H and $\delta^{13}C$ values of acetate even in pure standards. While there appears to be no correlation between acetate's ionization efficiency and its isotope ratios, there is a lower limit set by the magnitude and isotope composition of the background acetate signal. In Chapter 6, we found that samples of cow rumen
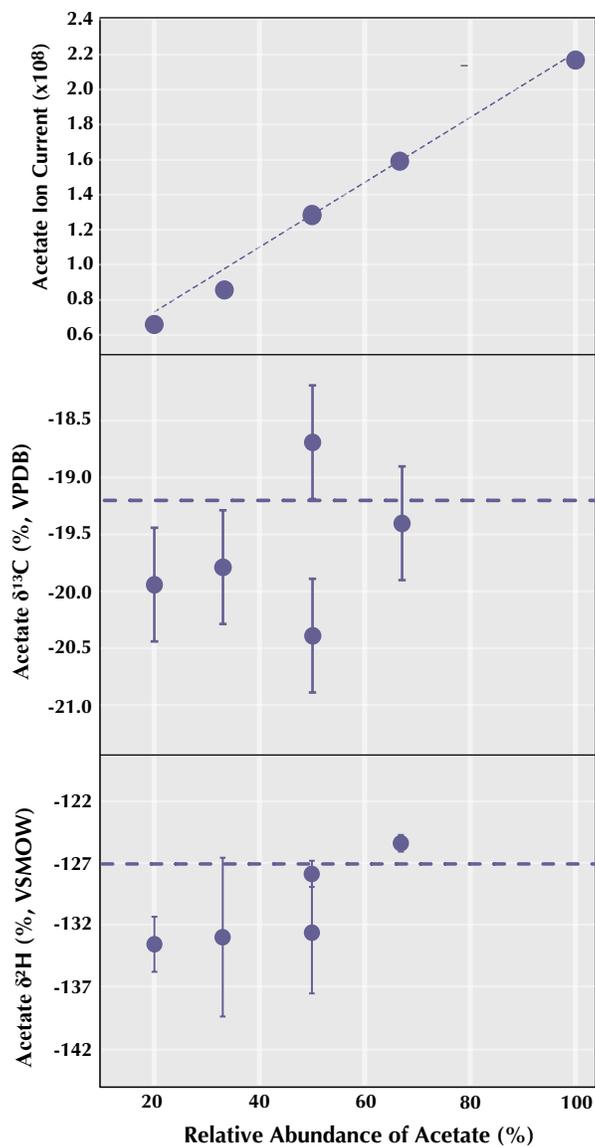
Figure 7.5: Propionate and butyrate cause ion suppression of acetate in the ESI source. When acetate's concentration was held constant in different organic acid mixtures (acetate + propionate + butyrate) and its relative abundance was decreased from 100% to 20%, its ion current decreased three-fold. This ion suppression did not cause demonstrable changes in acetate's $\delta^{13}C$ and δ^2H values.

where there were high concentrations of butyrate and propionate (72 hours, TMR+ Conditions, Supplementary Table 6.5), acetate had an anomalously positive $\delta^{13}C$ value (-12‰) compared to the rest of the student in which acetate drifted between -18 and -24‰. This fractionation was consistent between all incubation replicates of total mean ration (TMR) with seaweed additives. We attribute this fractionation to the lower ion signal (4×10^7) of acetate from the samples. The acetate suppression was caused by the sample's high butyrate-to-acetate ratio (red color, Figure 7.6).
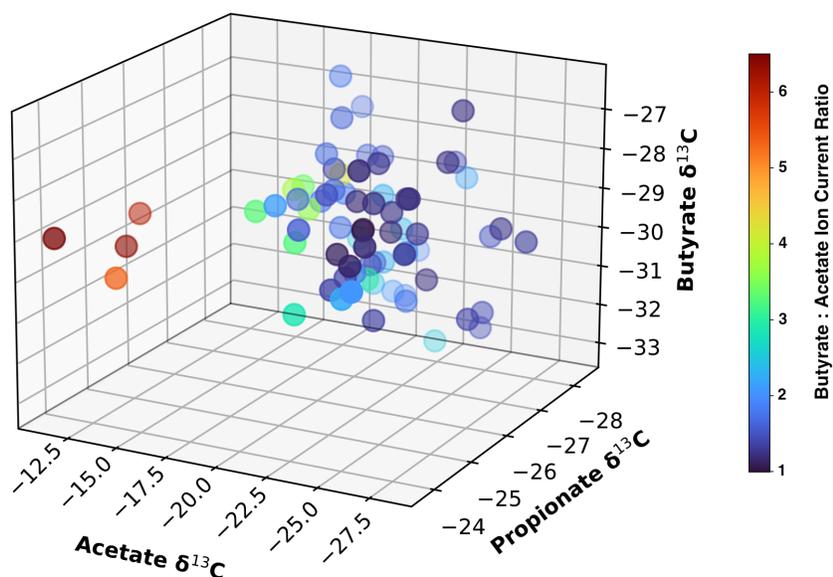
Figure 7.6: A three-dimensional rendering of compound-specific $\delta^{13}C$ values from acetate, propionate and butyrate in all samples measured in Chapter 6 from cow rumen. Colors indicate the ratio of butyrate and acetate's ion signals. Four red data points represent samples where acetate's ion intensity was extremely suppressed by butyrate. This in turn cased its measured isotope composition to become artificially enriched relative to other samples (blue and green). We attribute this to blank contributions that become more prominent when acetate's ion signal is severely suppressed ($<4\times10^7$).

Quadrupole Shifting

With multiple organic acids in the same solution, each of their isotopic properties are potential analytical targets. I developed a quadrupole "shifting" method that takes advantage of this opportunity. In Figure 7.7, a dual inlet introduction is used as an example of this technique. For the first half of a 16 minute acquisition, acetate

is measured by filtering for ions with a mass-to-charge ratio (m/z) between 57 and 62 (Figure 7.7A). At 8 minutes, the quadrupole shifts to a 72-75 m/z range, which captures propionate's isotopologues. After 5 minutes of measuring propionate's isotope composition, the quadrupole shifts again to analyze butyrate (Figure 7.7B). This sequence is then repeated for the next standard or sample (Figure 7.7C). At first, it may seem that quadrupole shifting is equivalent to the simpler approach of measuring one organic acid at a time. However, fewer total sample-standard switches are necessary. Since each switch requires about 10 minutes for autosampler injections, this saves 140 minutes per sample, assuming analytical triplicates are measured (Table 7.1, "Flushing Time"). It also decreases sample consumption since fewer injections are required, each of which wastes at least some sample (Table 7.1, "Waste Volume").

The timing of each quadrupole shift is intentionally asymmetric. Acetate ($C_2H_3O_2$), propionate ($C_3H_5O_2$) and butyrate ($C_4H_7O_2$) have progressively more carbon and hydrogen atoms. The relative abundance of their rare isotopologues increase proportionally. For example, acetate has two carbon atoms and the relative abundance of its ^{13}C-isotopologue is ~2.1%, while butyrate has four carbon atoms and a higher, ~4.2%, ^{13}C relative abundance. With more ^{13}C-isotopologues to count, butyrate requires fewer scans to achieve the same precision as acetate. In fact, the Orbitrap's mass resolution had to be increased from 60,000 to 120,000 in order to resolve the ^{13}C- and ^2H-isotopologues of butyrate (Table 6.1). This also decreased the scan rate of butyrate scans relative to acetate scans. However, with 2.5 minutes measurements, the precision on $\delta^{13}C$ and δ^2H values of butyrate match those of 8 minute acetate measurements. The total analysis time can be strategically allocated to maximize precision on each of the analytes of interest. This saves analysis time and decreases total sample consumption.
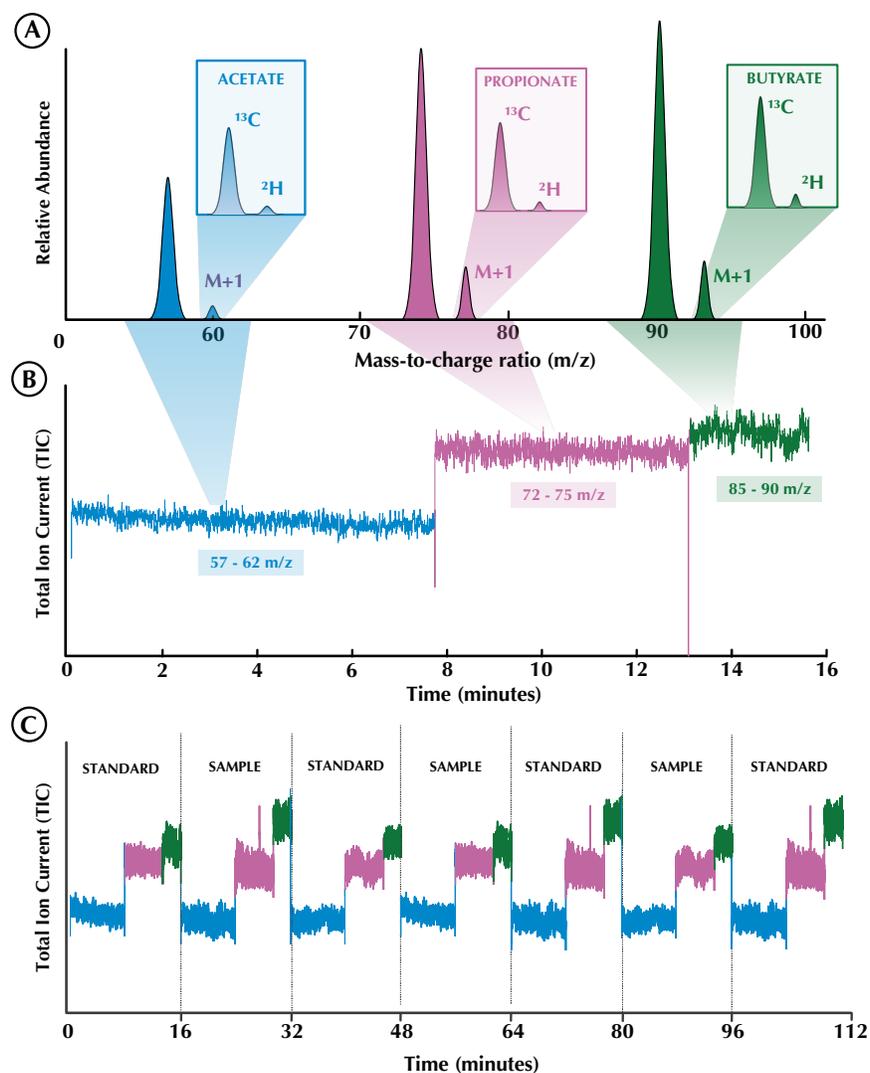
Figure 7.7: Schematic of the quadrupole "shifting" technique developed in this study to measure acetate, propionate and butyrate's $\delta^{13}C$ and δ^2H values in the same acquisition. A.) A representative mass spectrum if the quadrupole range was <50 to 100 m/z, with insets demonstrating the M+1 isotopologues. B.) Demonstration of quadrupole shifting. Within a 16 minute acquisition, time is allocated between three organic acids by shifting the quadruple between 57-62 m/z, 72-75 m/z, and 85-90 m/z mass windows. Asymmetrical allocation of time is intentional to maximize counting statistics on smaller molecules. C.) This sequence of quadrupole shifts is repeated for sample and standard to achieve triplicate $\delta^{13}C$ and δ^2H measurements. Ion signals for all three acids between sample and standard must be matched simultaneously.

An added complexity to the quadrupole shifting method is the need to sample-standard match across three analytes simultaneously. This is facile for samples with known concentrations, but for those with unknown concentrations (e.g. an unknown

yield from the isolation procedure), standards have to be carefully prepared that best match the sample. As seen in Figure 7.5, matching ion currents of sample and standards is not imperative for accurate measurements and there is a margin of error in matching sample and standard. This should be tested for each new analyte of interest with synthetic mixtures spiked with standards of known isotope composition. Using this technique, I successfully measured all 72 samples from the Chapter 6 rumen study with the autosampler introduction system.

Mass Spectral Contaminants

Contaminant ions that make it through the quadrupole filter can impact the measured isotope ratio of the anatlyical target. This has been shown previously in GC-Orbitrap methods (Hofmann et al., 2020). Hoffman et al. found that when contaminant ions appeared in the mass spectra of the sample, it can suppress the measured carbon isotope ratio. If this contaminant is not found in the standard at similar intensities, then this will cause a decrease in the measured $\delta^{13}C$ of the sample. For organic acid analyses on the ESI-Orbitrap, this phenomena appears as well. When measuring acetate, a common feature of the 57 - 62 m/z spectrum is a peak at 59.98 m/z which corresponds to a monovalent CO_3 ion (Figure 7.8A). It appears in the methanol blank and pure standards, though at low intensities typically <0.5% relative abundance.

In samples, the contaminant becomes more prominent for reasons that are still being elucidated. In one such sample from the rumen study, the 59.98 m/z ion increased steadily across three sample acquisition blocks within a dual inlet measurement (Figure 7.8B). There was an inverse correlation between the relative abundance of the carbonate ion and the measured $\delta^{13}C$ value of the sample. The three analytical replicates drifted down by over 20‰ over the course of the measurement. No ionization parameters were changed and the instrument was running uninterrupted for the entire measurement. The increase in carbonate ion likely came from a drift in the ESI source conditions over time. When the individual scans across these three sample acquisition blocks were concatenated and the carbon isotope ratio of acetate in these scans had an even more extreme relationship with the contaminant ion (Figure 7.8C). When the CO_3^- ion relative abundance was below 5-10%, the isotope ratio of acetate was consistent, but as the carbonate ion intensity increased, it drifted to smaller values, changing by as much as 50‰. Thus, the existence of contaminant peaks in the mass spectrum can suppress the isotope ratio of the target analyte. This is likely due to space-charge effects that preferentially interfere with the ^{13}C-isotopologue over the more abundant monoisotopic isotopologue (Hofmann

et al., 2020). In this case, 59.98 m/z is close to the ^{13}C-isotopologue of acetate (60.01 m/z), perhaps magnifying the effect.
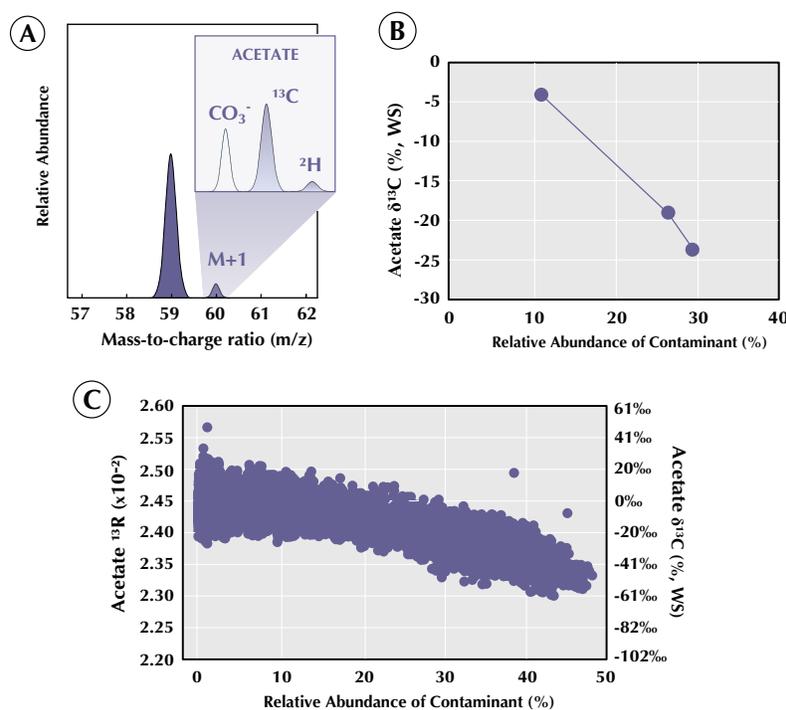
Figure 7.8: Contaminants within the mass spectra — close enough to target analyte masses that the quadrupole cannot filter them out — cause isotopic fractionations that are not replicated in standards, resulting in supresion of the observed $\delta^{13}C$ values. A.) Carbonate (CO_3^-) is adjacent to the ^{13}C-isotopologue of acetate in the mass spectra. B.) A single dual inlet measurement with three sample blocks. Each block had a different carbonate ion relative abundance, which correlated with the measured $\delta^{13}C$ values. C.) Collating individual scans across the three sample acquisition blocks, there was a strong correlation between contaminant relative abundance and acetate's $^{13}C/^{12}C$ isotope ratio. Moving average across 100 scans was plotted to smooth data. The isotope ratio was suppressed when the contaminant relative abundance exceeds 10%.

To mitigate this issue, carbonate ions had to be removed from the mass spectra. This was partially done in the sample prep, during which samples are titrated to a circumneutral pH, but not to a basic pH, which keeps CO_2 from dissolving more readily into the solution. The HPIC eluent is also dried down under nitrogen flow to degas the solution. However, the most effective strategy is within the ion source. Both decreasing the spray voltage and increasing sweep gas flow mitigated carbonate ionization. Unfortunately, the sweep gas flow also decreased the ionization yield of acetate by up to 50% in both samples and standards. In a trade off between

ion intensity and mass spectral purity, the latter is preferred until the ion intensity reaches low enough that blank contributions become an issue, as discussed above.

Conclusion

Bioanalytical mass spectrometry (e.g. ESI-Orbitrap MS) has greatly expanded the scope of molecular targets available for stable isotope ratio measurements at their natural abundances. To utilize these tools in environmental samples requires a commensurate effort in understanding matrix effects within ESI and APCI sources and their influence on measurement accuracy. In the first demonstrations of isotope ratio analysis on organic molecules from the environment using bioanalytical MS, we encountered such matrix effects. In response, we developed a workflow to purify our target analyte, acetate, from the environment while minimizing sample consumption and maximizing analytical throughput. We further identified mass interferences within the mass spectra that can worsen measurement accuracy. These studies represent a framework for identifying and mitigating problems as we transition this new technology from pure standards to natural samples. This framework is not comprehensive and will need to reevaluated and adapted for new analytes. However, it provides a roadmap for future work on organic analytes in nature.

Technique	Total Analysis Time	Flushing Time	Analysis Volume	Waste Volume	Total Volume
Direct Infusion	70 min	~2 min	150 μL	60 μL	210 μL
Dual Inlet	98 min	5 min	150 μL	340 μL	490 μL
Autosampler	140 min	10 min	150 μL	30 μL	180 μL

Table 7.1: Comparison of three sample introduction systems for measuring $\delta^{13}C$ and $\delta^2 H$ values of organic acids using ESI-Orbitrap. These estimates assume we are making three replicate sample acquisitions with sample-standard bracketing, totaling seven acquisitions. Each acquisition is assumed to take 10 minutes for actual analysis time, with a minimum 70 minutes of analysis time.

Technique	Automation	Carry Over	Flow Stability
Direct Infusion	X	XX	XX
Dual Inlet	XX	X	X
Autosampler	XXX	XXX	XXX

Table 7.2: Qualitative assessment of three sample introduction systems, X = sufficient, XX = good, XXX = great

References

Beulig, F. et al. (Jan. 2018). "Control on rate and pathway of anaerobic organic carbon degradation in the seabed". en. In: *Proc. Natl. Acad. Sci. U.S.A.* 115.2, pp. 367–372. DOI: 10.1073/pnas.1715789115.

Chimiak, L. et al. (Jan. 2021). "Carbon isotope evidence for the substrates and mechanisms of prebiotic synthesis in the early solar system". en. In: *Geochimica et Cosmochimica Acta* 292, pp. 188–202. DOI: 10.1016/j.gca.2020.09.026.

Csernica, Timothy, Surjyendu Bhattacharjee, and John Eiler (Aug. 2023). "Accuracy and precision of ESI-Orbitrap-IRMS observations of hours to tens of hours via reservoir injection". en. In: *International Journal of Mass Spectrometry* 490, p. 117084. DOI: 10.1016/j.ijms.2023.117084.

Csernica, Timothy, Alex L. Sessions, and John M. Eiler (Dec. 2023). "High-dimensional isotomics, part 2: Observations of over 100 constraints on methionine's isotome". en. In: *Chemical Geology* 642, p. 121771. DOI: 10.1016/j.chemgeo.2023.121771.

Eiler, John et al. (Nov. 2017). "Analysis of molecular isotopic structures at high precision and accuracy by Orbitrap mass spectrometry". en. In: *International Journal of Mass Spectrometry* 422, pp. 126–142. DOI: 10.1016/j.ijms.2017.10.002.

Heuer, Verena B. et al. (Dec. 2020). "Temperature limits to deep subseafloor life in the Nankai Trough subduction zone". en. In: *Science* 370.6521, pp. 1230–1234. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.abd7934. URL: https://www.science.org/doi/10.1126/science.abd7934 (visited on 04/19/2024).

Hilkert, Andreas et al. (July 2021). "Exploring the Potential of Electrospray-Orbitrap for Stable Isotope Analysis Using Nitrate as a Model". en. In: *Anal. Chem.* 93.26, pp. 9139–9148. DOI: 10.1021/acs.analchem.1c00944.

Hofmann, Amy E. et al. (Nov. 2020). "Using Orbitrap mass spectrometry to assess the isotopic compositions of individual compounds in mixtures". en. In: *International Journal of Mass Spectrometry* 457, p. 116410. DOI: 10.1016/j.ijms.2020.116410.

Kew, William et al. (Nov. 2023). "Natural Abundance Isotope Ratio Measurements of Organic Molecules Using 21 T FTICR MS". en. In: *Anal. Chem.* 95.47, pp. 17203–17211. DOI: 10.1021/acs.analchem.3c01816.

Liesack, Werner, Sylvia Schnell, and Niels Peter Revsbech (Dec. 2000). "Microbiology of flooded rice paddies". en. In: *FEMS Microbiology Reviews* 24.5, pp. 625–645. ISSN: 1574-6976. DOI: 10.1111/j.1574-6976.2000.tb00563.x. URL: https://academic.oup.com/femsre/article-lookup/doi/10.1111/j.1574-6976.2000.tb00563.x (visited on 04/19/2024).

Mueller, Elliott P., Alex L. Sessions, et al. (Jan. 2022). "Simultaneous, High-Precision Measurements of $\delta\ ^2$H and $\delta\ ^{13}$C in Nanomole Quantities of Acetate Using Electrospray Ionization-Quadrupole-Orbitrap Mass Spectrometry". en. In: *Anal. Chem.* 94.2, pp. 1092–1100. DOI: `10.1021/acs.analchem.1c04141`.

Mueller, Elliott P., Fenfang Wu, and Alex L. Sessions (Nov. 2022). "Quantifying Isotopologue Reaction Networks (QIRN): A modelling tool for predicting stable isotope fractionations in complex networks". en. In: *Chemical Geology* 610, p. 121098. DOI: `10.1016/j.chemgeo.2022.121098`.

Neubauer, Cajetan, Antoine Crémière, et al. (Feb. 2020). "Stable Isotope Analysis of Intact Oxyanions Using Electrospray Quadrupole-Orbitrap Mass Spectrometry". en. In: *Anal. Chem.* 92.4, pp. 3077–3085. DOI: `10.1021/acs.analchem.9b04486`.

Neubauer, Cajetan, Michael J. Sweredoski, et al. (Nov. 2018). "Scanning the isotopic structure of molecules by tandem mass spectrometry". en. In: *International Journal of Mass Spectrometry* 434, pp. 276–286. DOI: `10.1016/j.ijms.2018.08.001`.

Sherwood Lollar, B. et al. (Feb. 2021). "A window into the abiotic carbon cycle – Acetate and formate in fracture waters in 2.7 billion year-old host rocks of the Canadian Shield". en. In: *Geochimica et Cosmochimica Acta* 294, pp. 295–314. DOI: `10.1016/j.gca.2020.11.026`.

Thomas, Randal B (2008). "Intramolecular isotopic variation in acetate in sediments and wetland soils." en. In.

Ungerfeld, Emilio M. (Feb. 2015). "Shifts in metabolic hydrogen sinks in the methanogenesis-inhibited ruminal fermentation: a meta-analysis". en. In: *Front. Microbiol.* 6. DOI: `10.3389/fmicb.2015.00037`.

Wang, Zhenfei et al. (Mar. 2024). "Oxygen Isotope Analysis of Nanomole Phosphate Using PO_3^- Fragment in ESI-Orbitrap-MS". en. In: *Anal. Chem.* 96.11, pp. 4369–4376. DOI: `10.1021/acs.analchem.3c03070`.

Weiss, Gabriella M. et al. (Nov. 2023). "Analysis of intramolecular carbon isotope distributions in alanine by electrospray ionization Orbitrap mass spectrometry". en. In: *International Journal of Mass Spectrometry* 493, p. 117128. DOI: `10.1016/j.ijms.2023.117128`.

Wilkes, Elise B. and Ann Pearson (Nov. 2019). "A general model for carbon isotopes in red-lineage phytoplankton: Interplay between unidirectional processes and fractionation by RubisCO". en. In: *Geochimica et Cosmochimica Acta* 265, pp. 163–181. DOI: `10.1016/j.gca.2019.08.043`.

Zeichner, Sarah S., José C. Aponte, et al. (Dec. 2023). "Polycyclic aromatic hydrocarbons in samples of Ryugu formed in the interstellar medium". en. In: *Science* 382.6677, pp. 1411–1416. DOI: `10.1126/science.adg6304`.

Zeichner, Sarah S., Elise B. Wilkes, et al. (July 2022). "Methods and limitations of stable isotope measurements via direct elution of chromatographic peaks using gas chromotography-Orbitrap mass spectrometry". en. In: *International Journal of Mass Spectrometry* 477, p. 116848. DOI: `10.1016/j.ijms.2022.116848`.

A p p e n d i x A

QIRN USER GUIDE

Elliott P. Mueller[1]

[1]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA

Introduction

Quantifying Isotopologue Reaction Networks (QIRN) is a forward, numerical modelling tool that constructs reaction networks of any kind and reports isotopic compositions of substrates within that network. In this user guide, we will show you: 1.) How to generate the necessary input files for QIRN and 2.) How to use the associated graphic user interface (GUI).
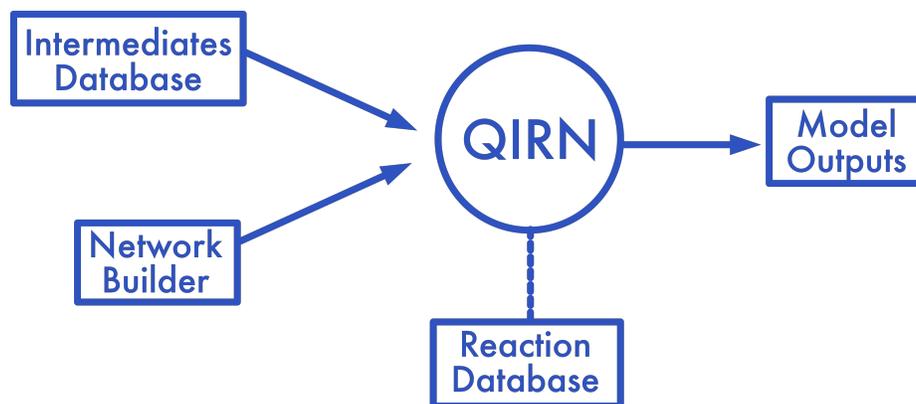
Figure A.1: The structure of how QIRN builds user-defined reaction networks.

Input Files

QIRN has three comma-separated-value (csv) files that it pulls from to construct user-defined networks. Two of these ('Intermediates Database' and 'Network Builder') the user interacts with. The final input ('Reaction Database') is the same for all users and is not changed, only referenced, by the user.

Intermediates Database

The Intermediates Database holds information about all the substrates that have ever been modelled in QIRN for a given isotope system. Here, the example is for carbon. This information includes the number of carbon atoms in the molecule, its starting concentration, its intramolecular isotope composition, and its boundary conditions. The atomic number never changes. The other three terms are changed by the user to define the initial conditions and boundary conditions of the model. If the "Reservoir" column is changed to "1", QIRN will keep the molecule's concentration and isotope composition constant over time. In the "Intramolecular" column, a vector of in numbers in delta-values can be input in the order of the molecule's atomic ordering. For example, glucose, can have a -10‰,-5‰,-10‰,-5‰,-10‰,-5‰ intramolecular structure from C1 to C6, respectively. These would be input as "-10,-5,-10,-5,-10,-5". If you are editing the csv file in Excel, the cell must be a "text" format.

IntermediatesDatabase

Intermediate	Number of Carbons	Initial Concentration	Intramolecular d13C	Reservoir
1,3-bisphosphoglycerate	3	0		
1,3-bisphosphoglycerate_cytoplasm	3	0		
2-isopropylmalate	7	0		
2-oxobutanoate	4	0		
2-phosphoglycerate	3	0		
3-4-hydroxyphenylpyruvate	9	0		

Figure A.2: Example of part of the Intermediates Database.

Reaction Database

The Reaction Database serves as a list of all reactions that QIRN has ever modelled and all their relevant information. Reactions are catalogued based on their Enzyme Commision (EC) number. If the reaction is not enzymatic a new ID may be given to it (i.e. 'TSPsulfate' for the transport of sulfate into a cell). If there is only one reactant in the reaction, only fill out the column for Reactant A. Similarly, only fill out the column for Product C if only one product is formed. In the case of a condensation function, QIRN will condense Reactants A and B (or C and D in the reverse direction) so that the last atom on A is connected to the first atom on B. For cleavage reactions of Reactant A to Products C and D, QIRN will cleave Reactant A based on the carbon numbering of Product C, which is reads from the Intermediates Database. Thus if Reactant A and Product C are six and four atoms long, respectively, Product C will be composed of Reactant A's sites 1-4 and Product D will be composed of Reactant A's sites 5 and 6. This can be changed by assigning a "Reaction Barcode" to the reaction. These strings of integers tell QIRN how to transition atoms from reactant to product and are applied to any type of reaction where atomic rearrangements occur. The Reaction Database must be in the same directory as QIRN's python scripts. This is a default when you download the QIRN installation package.

ID	Enzyme	Reactant A	Reactant B	Product C	P
EC 5.3.1.1	Triose phosphate isomerase	DHAP		glyceraldehyde-3-phosphate	
EC 5.3.1.1 Cytoplasm	Triose phosphate isomerase cytoplasm	DHAP_cytoplasm		glyceraldehyde-3-phosphate_cytoplasm	
EC 5.3.1.6	Isomerase	ribulose-5-phosphate		ribose-5-phosphate	
EC 5.3.1.9	Phosphohexose isomerase	glucose-6-phosphate		fructose-6-phosphate	
EC 5.4.2.11	Phosphoglycerate mutase	3-phosphoglycerate		2-phosphoglycerate	
EC 6.2.1.4	Succinyl-CoA Synthetase	succinylCoA		succinate	
EC 6.4.1.1	Pyruvate carboxylase	pyruvate	HCO3	malate	
EC 6.4.1.2	acetylCoA carboxylase	acetylCoA	co2_ACCase	malonylCoA	
F6PS	F6P Aldolase Sink	fructose-6-phosphate_cytoplasm_aldolase		glucose-6-phosphate	
F6PSII	F6P Aldolase Sink II	fructose-6-phosphate_cytoplasm		glucose-6-phosphate	
Lipid synthesis	Lipid synthesis	acetylCoA		Lipids	

Figure A.3: Example of part of the Reaction Database. "Product D" and "Reaction Barcode" columns are not shown here.

Network Builder

The Network Builder file is where users define the topology of their reaction network inlcuding, the reactions, their forward and reverse rate constants, and their isotope effects. Reactions are input as their EC number or other identifier defined by the Reaction Database. QIRN searches the Reaction Database for these reactions and will alert the user if any reactions do not match the database. QIRN then uses the Intermediates Database to select the appropriate substrates in the network, their starting concentrations and isotopoic compositions. Forward and reverse reaction rates refer to the mass action rate constant for monoiosotopic isotopologue. When defining rate constants, remember that steady state fluxes are a function of both the steady state concentration of the reactant(s) and the rate constant. Therefore, there is no simple relationship between the rate constant and absolute flux, unless all of the substrates are given a concenration of unity at the start of the simulation and rate constants at each substrate node are chosen so the inputs and outputs are equivalent. In this case, substrate concentrations are always unity and are constant and the system begins in steady state. Thus, the rate constants are equivalent to the steady state flux. This is the recommended strategy for choosing initial conditions in QIRN. For a discussion of exceptions to this, see the main text.

NetworkBuilder_glycolysis

Reaction	Forward Rate Constant	Reverse Rate Constant	Carbon Site of Fractionations Forward Reaction	KIE Forward	Carbon Site of Fractionations Reverse Reaction	KIE Reverse
EC 2.7.1.1	0.1	0				
EC 5.3.1.9	0.1	0				
EC 2.7.1.11	0.1	0				
EC 4.1.2.13	0.1	0				
EC 2.7.2.3	0.1	0				
EC 5.3.1.1	0.1	0				
EC 4.2.1.11	0.1	0				
EC 5.4.2.11	0.1	0				
EC 2.7.1.40	0.1	0				
EC 1.2.1.12	0.1	0				
EC 1.2.4.1	0.1	0	1,2,3	0.98,0.98,0.997		
EC 6.4.1.1	0.1	0				

Figure A.4: Example of a Network Builder file for the glycolysis pathway.

Site-specific kinetic isotope effects (KIEs) for both forward and reverse are also defined in the Network Builder. They are assigned as fractionation factors, where values less than one represent normal isotope effects (fractionation against the rare isotope). Equilibrium isotope effects (EIEs) are equivalent to the ratio of the forward to reverse KIEs. For example, if a reaction has no KIE but a 0.995 EIE, the user would input 1.0 for the forward KIE and 1.005025 for the reverse KIE. Each one of these KIEs is site-specific and must be assigned to an atomic site on the reactant to be imposed. This is assigned in columns 4 and 6 of the Network Builder file. Here, a list of numbers indicates the atomic sites where isotope effects should be imposed. The string of KIEs in columns 5 and 7 should be the same lengths as those in columns 4 and 6. For example, if the "Carbon Site of Forward Reaction Fowrard Reaction" is "1,2,3" and "KIE Forward" for that reaction is "0.98,0.98,0.997", then QIRN will assign 20‰ isotope effects on the C-1 and C-2 sites of the reactant and a 3‰ isotope effect on the C-3 site. In condensation reactions, the atomic sites and their KIEs are assigned from the first atom of Reactant A to the last atom of Reactant B. For example, if Reactant A and B are both C_2 molecules with isotope effects on the C-1 site of each, the user would input "1,3" in column 4 for that reaction. For the reverse reaction, the same principle applies, but the atomic sites are now structured from the first atom of Product C to the last atom of Product D.

Graphic User Interface
Running a Simulation

The graphic user interface (GUI) for QIRN gives users the ability to rapidly construct and anaylze isotopologue networks. The GUI requires two inputs from the user, the Network Builder and Intermediate Database csv files. After loading the GUI,

upload these files into QIRN.

Figure A.5: Opening page for the QIRN GUI where you can select the necessary input files from your computer.

They can be stored anywhere on your computer. To run a simulation, click "Run Forward Model". QIRN will ask you to input a length of time to run the model. This input is in seconds with timesteps occuring every 0.1 seconds. Thus, if you input 300 seconds, QIRN will run the model for 3000 timesteps. Each network will require different amounts of time to reach steady state, as discussed in the main text. We recommend starting with 100 seconds and adjusting from there based on the outputs. While the model is running, QIRN will display a progress bar with an estimated time to completion in the command window.

Figure A.6: Screenshot of the progress bar from a QIRN simulation

Output Visualization

Once the model has completed running, QIRN will provide four modules for visualizing the outputs. The first two are the concentration of substrates and fluxes of reactions. These can modules report each of these parameters in time over the course of the entire simulation. They are useful for ensuring that QIRN has created the correct network topology and for assessing whether the simulation reached steady

state conditions. The other two modules are for site-specific and compound-specific isotope compositions of each substrate. In the "Isotope Distributions" module, you can select individual substrates and QIRN will report their site-specific isotope composition. In the compound-specific module, QIRN will report the molecular average isotope composition of each substrate over the course of the entire simulation. This is a useful exercise for determining whether the model has reached isotopic steady state.
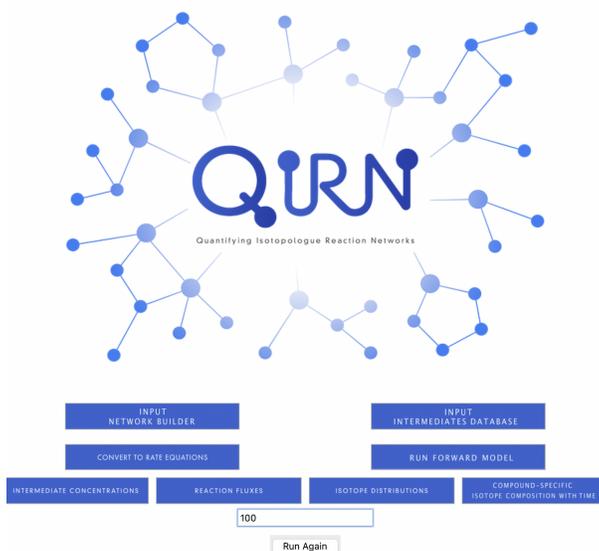
Figure A.7: Options for data visualization after a simulation has completed running.

Implementing Network Changes

After running a simulation, changes to network topology, reaction rate constants, isotope effects, boundary conditions or can be made easily. Simply make the necessary changes to the Network Builder or Intermediate Database which is already uploaded into QIRN, save them, and hit the "Run Again" button. There is no need to re-upload them into QIRN again. You can also replace the Network Builder without replacing the Intermediates Database, or vice versa. Only when changes are made to the Reaction Database does the GUI need to be reloaded.

Converting to Rate Equations

In simulations where substrates concentrations do not begin at steady state, the steady state fluxes of reactions do not have a simple relationship to the reaction rate constants. Therefore, we developed an optimization function that inverts rate

constants for reactions based on input steady state fluxes. To run this inversion, input a Network Builder and Intermediates Dtabase file for the network. In the rate constants columns of the Network Builder file input the requested steady state fluxes for each reaction. Put in a specified time in which you would like to run the network to steady state. The rate constants inverted will depend on this time. If QIRN finds a set of rate constants that reaches the steady state fluxes, it will create a new csv file labelled with the name of your Network Builder file with an attached label of "FLUXINVERTED". This file is identical to the input Network Builder file with the inverted rate constants filled into the forward and reverse rate constant columns. In the command window, QIRN will report the residual between the requested steady state fluxes of reactions and those which are produced from the optimized reaction rate constants.