CREATING A FRACTION LIBRARY OF THE *THALASSIOSIRA PSEUDONANA* (CCMP1335) ENDOMETABOLOME

by

CIONDI FELICIANEAU BESS

(Under the Direction of Arthur S. Edison)

ABSTRACT

Marine phytoplankton fix carbon dioxide during photosynthesis and contribute half of the carbon fixation occurring on Earth. During photosynthesis, small organic molecules are produced within the cell (i.e. endometabolites) and released into the seawater to support heterotrophic organisms in the ocean. Improving our knowledge on phytoplankton endometabolites can bring deeper understanding of phytoplankton metabolism and how it effects the global network of life. The universal problem arises with identifying metabolites from complex mixtures of compounds. Here, we take an approach of making a fraction library where metabolites were separated into fractions by polarity. A common phytoplankton strain *Thalassiosira pseudonana* (CCMP1335) was used as a representative model for this analysis. Nuclear magnetic resonance (NMR) spectroscopy and a recently developed database matching approach were used for compound annotation. Separating the endometabolome into fractions improves compound identification and supports comprehensive functional analyses that bring us closer to metabolite identification.

INDEX WORDS: Endometabolome, Phytoplankton, Metabolites, Nuclear magnetic resonance spectroscopy, Chromatography, Fraction library, *Thalassiosira pseudonana*, Metabolomics, Spectral Annotation by Feature Extraction and Reference matching

CREATING A FRACTION LIBRARY OF THE *THALASSIOSIRA PSEUDONANA* (CCMP1335) ENDOMETABOLOME

by

CIONDI FELICIANEAU BESS

BS, SOUTH CAROLINA STATE UNIVERSITY, 2020

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2024

CREATING A FRACTION LIBRARY OF THE *THALASSIOSIRA PSEUDONANA* (CCMP1335) ENDOMETABOLOME

by

CIONDI FELICIANEAU BESS

Major Professor:	Arthur S. Edison
Committee:	Erin Dolan
	Mary Ann Moran
	Michael Tiemeyer

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
December 2024

DEDICATION

To my Family, thank you for your many expressions of support. From financial contributions, prayers, words of encouragement, phone calls, etc., you have made it possible for me to push through, even beyond this day. I do it for all of you!

To the literal love of my life, Bruce, thank you for being there when I needed to lean. The support and love you have shown from the day I got accepted into the program has shown me that you will be here through it all! Thank you for pushing me, holding me accountable, and keeping me focused on the prize, despite the obstacles in the way. No matter how hard it got, you always reminded me where my help comes from. You're my biggest supporter!

To 'Molecules of Melanin', Friends and Associates, thank you for allowing our spaces to be safe. This space was REQUIRED to stay grounded and to remain motivated while transitioning and being in such an uncomfortable space. I am grateful to have experienced such great people.

To all Scientists, Chemist, Biochemist, Biologist, that may feel imposter syndrome at any point, those who are unhappy with where they are in life, and to those that are unhealthily overachieving for praise, please take a step back and just reflect. Feel your feelings. Deal with your emotions. Understand that success will still be there, and your happiness is way more important that the degree you're working on or the job you feel stuck in. Be intentional in everything that you do.

To Thee South Carolina State University, I hope that I continue to make you proud as an active Alum! I'll Defend and Honor, Love and Cherish Thee.

ACKNOWLEDGEMENTS

I would like to thank my committee Dr. Arthur Edison, Dr. Erin Dolan, Dr. Mary Ann Moran, and Dr. Michael Tiemeyer. A super-duper big thank you to Art for the support, trust, encouragement, patience, and guidance in every encounter we've had. I also want to thank my mentor, Mario Uchimiya, and colleague, Hope Cobbs, as well as the undergraduate researchers (VIPs): Kennedi Clemons, Kara Garrison, Cassandra Leahy, and Abby Modinski, for being such great researchers and trusting my guidance. Thank you to Chris Esselman and Conrad Epps for always answering a million questions and helping me in any way. Lastly, Thank you to everyone apart of the Edison Lab for your support

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

1.1. Carbon Cycling in The Ocean

In the surface ocean, half of the Earth's carbon is fixed by phytoplankton during photosynthesis (Falkowski et al., 1998; Holderman et al., 2023). Phytoplankton produce half of the oxygen on planet Earth (Moran et al., 2022), which make them essential to marine and human life. About a half of this fragment is released from phytoplankton into the surrounding seawater as dissolved organic matter (DOM) (Moran et al., 2022). Among the released DOM, metabolites that are rapidly (hours to days) used by other heterotrophic bacteria (i.e. labile metabolites) in sea water contribute to the dominant portion of carbon flux within the system (Moran et al., 2016). Heterotrophic bacteria depend on the flux of labile DOM to thrive (Holderman et al., 2023; Moran et al., 2022; Moran et al., 2016). The release of DOM is rapidly consumed up by surrounding marine organisms and used as nutrients (Moran et al., 2022). As the Earth heats up, the ocean stores more CO_2 than in the global atmosphere (Siegel et al., 2023). Excessive storage of CO_2 can result in its release back into the atmosphere and acidification of the ocean (Falkowski et al., 1998; Holderman et al., 2023). The effects of climate change and other ecological factors on phytoplankton in the marine environment cause ripple effects on the marine ecosystem (Moran et al., 2022).

The National Science Foundation (NSF)-funded grant for the Center for Chemical Currencies of a Microbial Planet (C-CoMP) focuses on the microbial role in the air-sea exchange of carbon. There is limited information on the chemical pathways and syntheses of organic and inorganic compounds by marine phytoplankton and the mechanisms to excrete these metabolites. This

problem originates from two factors. Firstly, it is difficult to measure the concentrations of released metabolites (i.e. exometabolites) in the surface ocean, because most metabolites are taken up just as fast as they are released from phytoplankton (Moran et al., 2022). Secondly, measuring metabolites in salty samples poses a challenges and limitations in the efficiency of analytical techniques (Edison et al., 2021; Moran et al., 2022).

1.2. Phytoplankton and Their Relationship with Heterotrophic Bacteria

Phytoplankton are microscopic marine organisms that are just as essential to aquatic life as they are to human life. They are the primary source of metabolites for surface ocean bacteria and are responsible for providing half of Earth's oxygen (Ferrer-Gonzalez et al., 2021; Moran et al., 2022). Phytoplankton are about 3 - 200 μm in size but have a huge impact in marine food webs. Carbon cycling is indispensable without phytoplankton. Changes in concentration of CO_2 are predicted to cause phytoplankton metabolism to be modified for survival (Clement et al., 2017). Phytoplankton provide carbon and nutrients by photosynthesis of CO_2 and through excreting, leaking, lysing, and being grazed (Durham et al., 2015; Moran et al., 2022). They actively and passively release, and in some cases, prematurely leak metabolites into the surrounding sea water, depending on the changes in their environment (Seymour et al., 2017). Metabolites that have been released from the cells and are present in the surrounding water are called exometabolites. Exometabolites are consumed by neighboring heterotrophic bacteria and can sometimes be transported to deeper parts of the ocean (Moran et al., 2022; Siegel et al., 2023).

Phytoplankton and heterotrophic bacteria both supply nutrients to the other for survival (Ferrer-Gonzalez et al., 2021; Seymour et al., 2017). The exometabolites that heterotrophic bacteria rapidly consume from phytoplankton make them less available for analysis. The lack of

knowledge of the exometabolite pool is a major focus for marine scientists as we learn more about the relationship between phytoplankton and heterotrophic bacteria.

Heterotrophic bacteria require nutrients from other organisms to stay alive and function properly. *Ruegeria pomeroyi* is a model heterotrophic bacterial species that has an affinity for the phytoplankton *Thalassiosira pseudonana*. In cocultures, *R. pomeroyi* relies on the exometabolites released extracellularly by *T. pseudonana.* Phytoplankton metabolite release depends on both the concentration of metabolites within the cell and physiological conditions of the cell (Uchimiya et al., 2022). It is believed that if the phytoplankton release of metabolites is altered in any way, the remaining organisms in the food chain are subject to changes in their heavily regulated systems as well. Therefore, phytoplankton's contribution to Earth is vital.

1.3. Endometabolome vs Exometabolome

Ways in which phytoplankton metabolites enter and exit the cell are still being investigated through research conducted to understand the biochemical currencies within phytoplankton cells. Metabolites that are within the cell are called endometabolites. The cell houses a million-fold higher concentration of metabolites compared to what is outside of the cell membrane (Moran et al., 2022). Being that exometabolites in seawater are more diluted than the endometabolome, it is more difficult to analyze them with analytical experiments. The first problem arises from the high level of salt in seawater affecting analytical techniques like liquid chromatography (LC), mass spectrometry (Jonker et al.), and nuclear magnetic resonance (NMR) spectrometry (Edison et al., 2021). The saltiness of the exometabolome sample can lead to inconsistent chromatography across identical high performance liquid chromatography (HPLC) injections and to reduced sensitivity in NMR spectrometry (Wishart et al., 2022).

To address the problem of dealing with a saltier exometabolome sample, a benzoyl chloride derivatization and solid phase extraction (SPE) steps would be helpful prior to liquid chromatography and NMR spectroscopy (Edison et al., 2021; Widner et al., 2021). The benzoyl chloride derivatization would consist of basifying the exometabolome sample with sodium hydroxide, adding benzoyl chloride, acidifying with phosphoric acid, then allowing gravity-elution on SPE resin to remove the remaining salt in the sample (Widner et al., 2021). These steps are beneficial to the exometabolome because it increases the sensitivity of the knowingly low concentration of metabolites and removes salt from the sample, but the derivatization also brings an additional problem. Adding benzoyl chloride will change the metabolic composition of the exometabolome, pushing us further from accurate compound identification. The derivatization also becomes laborious in sample preparation, when dealing with larger fraction sets (Zhang et al., 2014). The endometabolome does not require the benzoyl chloride derivatization and reduces experiment time, further supporting experimentation on the endometabolome.

The third problem faced, as mentioned above, is the low concentrations of metabolites for our analysis (Moran et al., 2022; Moran et al., 2016). The exometabolites are diluted by the seawater, making detection and annotation more difficult (Edison et al., 2021). Concentrating the endometabolome extraction sample will bypass issues in detection. With all things considered, it is expected that the endometabolome will provide detailed information about available compounds based on a higher concentration of metabolites. It is understood that these metabolites can be extracted from the endometabolome, yet the problem still arises as to what the exchange of metabolites between the endometabolome and the exometabolome are. Annotation and comparison efforts of both environments will increase the knowledge of phytoplankton metabolomics.

1.4. Challenges for Metabolomics

1.4.1. Nuclear Magnetic Resonance Spectroscopy vs Mass Spectrometry

Effort to create an accurate metabolome for phytoplankton have been delayed due to difficulties in concentration, chemical approaches and salt (Moran et al., 2022). There are numerous ways to begin a metabolic compound analysis, but they all have advantages and disadvantages. NMR spectroscopy and liquid chromatography - mass spectrometry (LC-MS) are the two most common analytical techniques used for metabolite analysis (Edison et al., 2021). NMR is a non-destructive, highly reproducible technique rich in details about the composition of a sample. Compounds are easily overlooked in chromatography without the use of NMR (Edison et al., 2021). NMR magnet size and range in experiments can vary based on research needs and this serves as a strength to the technique (Edison et al., 2021). NMR sensitivity improves at the use of a higher magnetic field and smaller probes (Edison et al., 2021; Wishart et al., 2022). Contrary to NMR, mass spectrometry is a destructive and more sensitive analytical technique that measures the mass to charge ratio of a set of compounds in a sample. MS can also assist in metabolite analysis with the support of other complementary analyses. The reproducibility in mass spectrometry and nuclear magnetic resonance help reduce variation in sample preparation. This becomes an advantage when experiments need to be repeated and compared.

Combining data from MS and NMR is a complicated process since there are numerous distinctive variables for both techniques. Some research efforts use MS and NMR as two individual approaches for metabolite analysis, whereas, in this analysis, they are used in complement of each other. In this set of experiments, a high-resolution LC/MS was not used. Instead, the mass spectrometer was connected to an HPLC system and served as the detector for the system. The chromatographic separation was collected into fractions and was analyzed by NMR. The

chromatographic MS signal shows how concentrated a fraction may be, as the measured abundance begins to reach its maximum. More concentrated fractions will consist of a higher abundance signal. Signals from a concentrated sample that have reached the maximum level of detection cause overlap and peak distortion, challenging the data quality of the analysis.

1.5. Fraction Library

The idea of separating chemical compounds from a complex biological sample would be useful for any scientific field. (Edison et al., 2021; Jonker et al., 2019). Metabolomics is the study of metabolism in an organism. In these experiments, phytoplankton metabolomics are used as a gateway to abundant information, bridged together for the understanding and identification of metabolites within phytoplankton cells. With the marine environment being rich in metabolites, analyzing in smaller portions would provide more detail. A library of fractions that eluted from liquid chromatography would serve as a great tool for the annotation of compounds and for additional biogeochemical research (Edison et al., 2021). A "Fraction Library", as it is called, is a spread of fractions that are collected from liquid chromatographic elution, analyzed, and stored for further scientific use. Liquid chromatography was incorporated to moderately separate compounds prior to assay application (Wilson et al., 2020). For this project, the endometabolome mixture was injected into the system and pushed through to the column with the mobile phase flowthrough. As the endometabolome mixture and mobile phase flowed through the column, compounds in the mixture adhered to the column resin based on their functional group binding properties. As the run progressed and mobile phases flowed at different concentrations, compounds will be gradually released from the packed column -- this is called elution. The direct relationship (increasing and decreasing at an equal rate) of both aqueous and organic phases over time is known as a gradient.

A silica-packed HPLC hydrophilic interaction (HILIC) column was used to separate the extracted endometabolome mixture by polarity. More polar compounds elute from the column later in the gradient while less polar compounds elute earlier. As the flow of the aqueous mobile phase increased (making the flow of the organic mobile phase decrease), the flowthrough become more polar and more polar compounds gradually eluted from the column. Based on the run time, 140 fractions were consecutively collected into vials. Separated fractions have been shown to improve assay performance and increase concentration (Wilson et al., 2020). Each individual fraction was analyzed by NMR. Overlap across fractions does occur depending on similar polarities of compounds or if there is an abundance of similar compounds eluting at similar retention times. The fraction library can serve purpose for further scientific research of how phytoplankton chemically transform and exchange metabolites with neighboring surface ocean marine life.

A fraction library will be a resourceful tool, containing an abundance of knowledge about the chemical compounds in a mixture (Edison et al., 2021). Low costs of phytoplankton culturing and extraction, while ensuring high reproducibility in sample preparation will improve the frequency at which a fraction library is used (Wilson et al., 2020). Separation of metabolites in a mixture is a significant part of the fraction library analysis, because without good separation, peaks will overlap, causing compounds in a lower concentration to remain hidden under highly concentrated compounds. Bad separation could also be a result of a column that does not fit the needs for separation, causing the contents of a sample to elute prematurely and simultaneously and not adhere to the column, despite the flow of the mobile phase gradient. To annotate the constantly changing metabolic state of the endometabolome, sufficient sample material, time, and methodical techniques are required to minimize variation that could lead to bias in experiments (Judge et al., 2019).

CHAPTER 2

METHODS

2.1. *Thalassiosira pseudonana* (CMP1335) Cell Culturing

The *T. pseudonana* (CMP 1335) cells were cultured using a similar culturing method as described in their publication (Olofsson et al., 2022).

Thalassiosira pseudonana CCMP 1335 phytoplankton (National Center for Marine Algae and Microbiota) were grown in L1 media supplemented with silica in glass bottles. It was important to refrain from exposing the media to light because the vitamins are photosensitive. The bottles were stored at 4°C and covered in tin foil to prevent any interaction with light. Prior to culturing, these glass bottles were acid washed for 3 hours, washed with deionized water, left to air dry, combusted at 450°C for 4 hours, and autoclaved for sterility. During culturing, vials tops were slightly lifted to allow proper air flow for the phytoplankton to grow. *T. pseudonana* (CMP1335) was inoculated at 1%. The media was changed on these cultures weekly (7 days). Sterility of the cultures was confirmed by plate check. On day 7, the *T. pseudonana* (CMP1335) culture was filtered using a 2.0-µm-isopore-filter (Millipore Sigma) by gentle vacuum filtration. The filter was then stored in 2 mL cryovial and transferred to a -80°C freezer to prevent degradation.

2.2. Endometabolome Extraction

In this study, a previously described MeOH extraction method was used (Holderman et al., 2023), with minor adjustments.

Three 2 mL cryovials containing a filter of 500 mL *T. pseudonana* (CCMP 1335) culture were removed from the -80°C freezer. Sterile forceps were used to remove the filters from the tubes, transferred into 50 mL conical tubes (Falcon, Corning Incorporated), and placed onto ice. Sonication was used to break up the phytoplankton membranes and to release the cells from the filter. Ice and water were added to the ultrasonic bath (Fisher Scientific FS-140H) in the event the cells began to heat up as they were broken by sonication. 15 mL of HPLC LC/MS-grade water was added to each filter in their conical Falcon tubes and placed inside the sonicator. The sonication switch was shifted on and off every 10 seconds, for a total time of 1 minute and 30 seconds for sonication. Cells were detached from the filters and displaced in the water. The water containing the lifted cells was transferred to three clean 50mL conical Falcon tubes, capped, and placed on ice. The sonication and washing steps were repeated two additional times, for a total volume of 45 mL in each conical Falcon tube containing the lifted cells and water. The filters remained in their initial conical Falcon tubes and placed in the -80°C freezer. The cells were stored in the -20°C freezer overnight to slow degradation. Once the extraction was frozen, it was transferred to the -80°C freezer to ensure that the sample was completely frozen, in preparation for lyophilization. The sample being completely frozen is vital to the endometabolome extraction because the lyophilizer uses constant vacuum pressure to not only maintain the frozen state of the sample and reduce degradation, but to also sublimate the solid form of the solvent within the vacuum system. Lyophilization is known for being consistent in results with minimal variation during processing (Holderman et al., 2023).

The lyophilizer (VirTis Benchtop™ “K”) was checked prior to removing the extraction from the -80°C freezer. A clean, glass lyophilization jar and lid was used to freeze dry the extraction. The extraction was removed from the -80°C freezer and placed onto dry ice to slow down the

potential thawing of the sample. As room temperature is very warm compared to -80°C, it will begin to thaw quickly if not placed on dry ice and its frozen state kept intact. The extraction needs to remain frozen for successful lyophilization. The cap of the conical Falcon tube containing the extraction was replaced with parafilm tape (ParaFilm Lab Film) securely placed over the mouth of the conical tube. Small holes were gently poked through the stretched parafilm with a sterile knife to ensure the sample could be under vacuum and freeze drying could occur. The conical Falcon tube was placed into the glass lyophilizer jar and attached to the lyophilizer before the vacuum valve was opened. Once the jar was secured, the valve was slowly turned to open, ensuring there was not a drastic pull in vacuum or a dramatic change in pressure by adding the jar. The lyophilization start time and stabilizing pressure were recorded. The jar stayed on the lyophilizer until it was completely dried and there was nothing frozen remaining in the tube.

The jar was checked periodically to ensure the extraction was drying and that the pressure on the lyophilizer was stable. Once dried, the vacuum valve was slowly closed, and the jar was removed. The jar was transferred to a 4°C cold room. The dried extraction was resuspended into 1 mL of HPLC LC/MS-grade water by gradually adding in 100 µL increments. The extraction mixture was vortexed for 10 minutes to ensure cells were broken to maximize the mass of endometabolome extraction. The dissolved extraction was centrifuged (Eppendorf Centrifuge 5417C) at 4°C for 10 minutes at 14,000 RPM. The supernatant was collected, leaving the pellet in the tube and aliquoted into 200 µL increments. The tubes were centrifuged for 30 seconds at 4,000 RPM to gather all contents of the endometabolome extraction to the bottom of the tube to evenly freeze when placed in the -20°C freezer. After the sample was frozen, it was moved to the -80°C freezer until future analysis was ready to be performed with the material.

2.3. High-Performance Liquid Chromatography (HPLC) Fractionation

Two 200 µL *T. pseudonana* extraction vials were removed from the -80°C freezer and placed on the vacuum centrifuge (Labconco, CentriVap Concentrator). The sample was stored in 100% LC/MS grade water, making the removal of any solvent within the extraction essential for resuspension and HPLC fractionation. As the sample dried, a pellet formed in the bottom of the tubes, containing the concentrated endometabolome. 150 µL of 80/20 Methanol/Water was added and vortexed to break up both extraction pellets, resuspending the contents of the endometabolome. Both samples were combined into a specialized HPLC vial [Thermo Scientific, Clear Glass Micro-V Tapered MicroVial 150 µL Reservoir 1.5 mL capacity] with a tapered bottom. Using a tapered bottom vial allowed any undissolved or nonpolar material to be collected and prevented touching by the injection needle on the HPLC system. The sample was placed in the 4°C refrigerator while the HPLC with LC/MS system was equilibrated. The system was manually washed at varying percentages of both mobile phases, without the column attached. This allowed both aqueous and organic mobile phases to flow through the system and remove any particulates that may have been left behind from previous uses or what could have accumulated while being inactive. Mobile phase A Durham et al. (2015) consisted of 0.1% formic acid in water and mobile phase B (MPB) consisted of 0.1% formic acid in acetonitrile. The gradient method included variations of gradually changing aqueous and organic mobiles phases over 35 minutes. Water, being the most polar compound, is electronegative neutral. The neutral charge allows for the molecule to bind to itself or other molecules, making it appropriate for this polarity-based separation. Acetonitrile (ACN) contains carbon and nitrogen, making its polarity lower than water. Both mobile phases contained formic acid to help with reducing ion suppression in MS. The gradient elution method was compared to other fraction library methods for other organisms within

the lab. It was important to cross check the gradient methods for fractionation to support the hypothesis that the fraction library is useful to different small organism's metabolome when using this direction to complete compound analysis. Within the method, the injection needle height was raised to 3 mm to ensure that the needle would not touch the undissolved material in the bottom of the vial.

In preparation for the column equilibration, an XBridge BEH Amide [OBD Prep Column, 130Å, 5 µm, 10 mm X 250 mm] HILIC column was connected to the system, ensuring both ends of the column were securely screwed into the connecting lines. This normal phase HILIC column provides polar separation chromatography of metabolites ranging from pH 2 to pH 11. The column was placed in the column heating compartment to improve separation. There is an arrow on the outside of the column that shows which way the flowthrough should go. The mobile phases flowed from the quaternary pump, through the column in the direction of the arrow. The eluent splits to the mass spectrometer, used as a detector, to produce a chromatograph and to the fraction collector for the collection of fractions. 95:5 MPA:MPB flowed through the column to check for leaks from either end of the column. Any column leakage would be visible, and the pressure of the system would be extremely high and unstable as the leak continues. After about 5 minutes of the equilibration at 95% MPA, the percentage gradually decreased to 5%, aligning with the method parameters for the elution gradient (**Table 1**). The column heater was set to 25°C to increase column separation, and the fraction chiller was set to 4°C to reduce degradation of metabolites as they were collected. Both compartments were allowed to reach the set temperature before proceeding with start the run.

Table 1 HPLC HILIC Gradient Elution Method.
HPLC system used a normal-phase HILIC column to separate the complex sample by polarity. Mobile Phase A consists of 0.1% Formic Acid in Water. Mobile Phase B consists of 0.1% Formic Acid in Acetonitrile.

Time [min]	MPA [%]	MPB [%]	Flow [mL/min]	Max Pressure Limit [bar]
00.00	5.0	95.0	3.500	550.00
20.00	30.0	70.0	3.500	550.00
30.00	50.0	50.0	3.500	550.00
35.00	65.0	35.0	2.000	550.00

This gradient elution uses a mass spectrometry detected signal to produce a chromatograph of the separated endometabolome as it elutes from the HILIC type column. The elution from the column splits between the mass spectrometer and the fraction collector. Within the system, there is a delay in the fraction collector, so the visual chromatography and fractions collected align, since the MS signals must be read first. This delay requires a "make-up" pump to equate for the volume destroyed in the mass spectrometer. Fractions are collected on an untargeted, time-course basis (i.e. a set number of fractions over a given time span).

The endometabolome sample was removed from the 4°C fridge and placed onto the HPLC system. Three 100 μL injections from the sample of endometabolome were performed. The total run time for the method was 35 minutes with 140 fractions collected in numerical order, from 2.00 minutes to 34.00 minutes (32 minutes of collection time). Since the mass spectrometer data was collected without a set of targeted compounds of interest, the collection of fractions was considered untargeted, time-course fractionation (**Figure 1**). As the organic mobile phase decreases and the aqueous increases at the same rate, more polar compounds will elute with the increase in aqueous mobile phase, since water is the most polar compound.

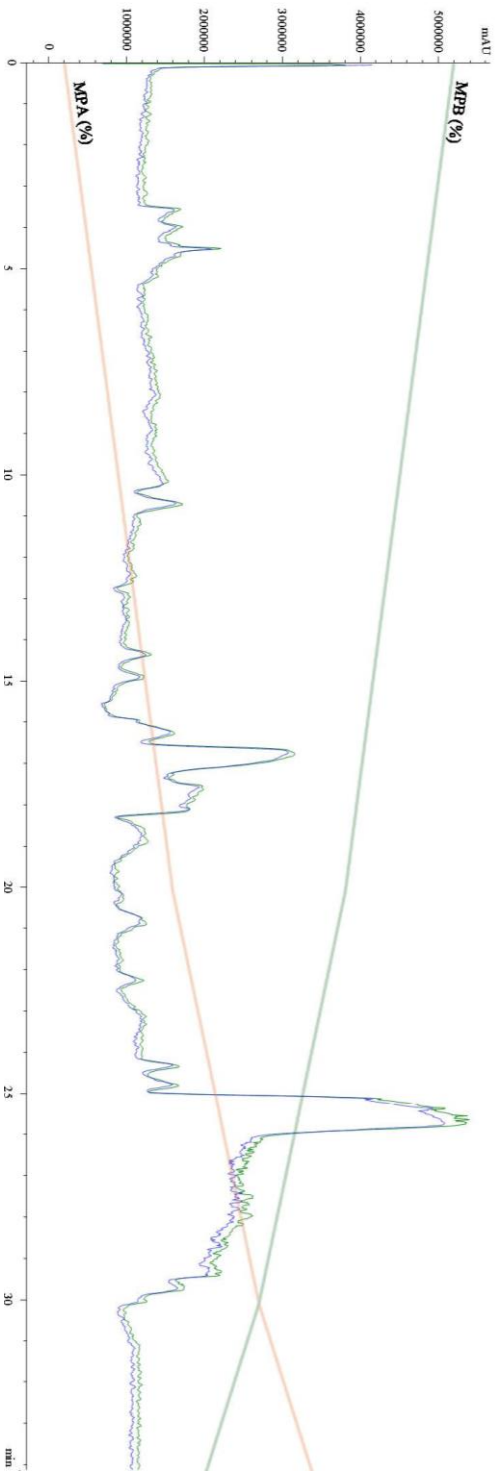

Figure 1 Two Injections of Fractionated Endometabolome. Two Identical Chromatographs from two HPLC injections. The blue and green chromatograms represent differences in injections. The orange line indicates the increase in Mobile Phase A (0.1% Formic Acid in Water) and the blue line indicates the decrease in Mobile Phase B (0.1% Formic Acid in Acetonitrile) during the gradient elution. This signal was mass detected using a total ion count on a single quadrupole mass spectrometer. 140 Fractions were collected from 2.00 minutes to 34.00 minutes.

After the third injection finished, 10% of each fraction's total volume (about 80 µL) was transferred into a different HPLC vial and set aside for potential HPLC mass spectrometry/ mass spectrometry (MS/MS) analysis to be performed in the future. The endometabolome fractions and the MS/MS fractions were dried in the vacuum centrifuge separately. All material was stored at -80°C until prepared for NMR data acquisition.

2.4. NMR Data Collection and Processing

The dried, fractionated endometabolome was removed from the -80°C freezer and thawed at room temperature. In preparation of the NMR data collection, $\frac{1}{9}$ mM DSS in D_2O + 100 mM sodium phosphate buffer was made. DSS is a common NMR reference solvent that has a chemical shift value at 0.0 ppm, regardless of the contents of the sample. D_2O is also a common NMR solvent that is responsible for locking the spectrometer to maintain a stable magnetic field around the D_2O signal, despite any shifts that may occur in the sample set. D_2O also reduces the water signal in NMR samples. Water shows an intense signal in proton (^1H) NMR and can overpower other compound signals. Suppressing this water signal is beneficial but is also responsible for signal loss from about 0.5 - 1.0 ppm of the water peak (Edison et al., 2021). With a smaller water peak, the baseline noise is reduced, causing smaller signals to be intensified and observed in relation to larger signal peaks belonging to their metabolite or other metabolites at similar chemical shifts. The sodium phosphate buffer added to the NMR sample, in addition to the reference material, helps with maintaining pH within the fractions and helping with consistent chemical shift values of compounds across separation. The phosphate buffer acts as a pH stabilizer across the fractionation. Even though the compounds eluted by polarity, the pH of the fractions can vary. 50 µL of the NMR solvent was added to all 140 fractions and vortexed to ensure the metabolites were

dissolved in the NMR solvent. Each fraction, now containing metabolites dissolved in 50 μL of NMR solvent, was transferred into a 1.7 mm Bruker BioSpin Sample Jet NMR tube. To confirm that the NMR solvent was not contaminated during the repetitive transfer from the solvent vial, a "blank" containing only the NMR solvent was taken after every 35th fraction (not shown in **Figure 2**). The blanks were added to the NMR cassette as they were taken, causing the tube rack numbering to be slightly shifted. There was a total of 144 tubes, including the 140 fractions and 4 blanks across two cassette racks.

Both cassette racks were transferred to the autosampler attached to the magnet in the Nuclear Magnetic Resonance Facility, located at the University of Georgia. The magnet used was an 800 MHz Bruker Avance NEO NMR with a 1.7 mm TCI cryoprobe. The cooling and use of a cryoprobe reduced electronic noise in the spectra, improving the signal to noise ratio (Edison et al., 2021). Before the data could be acquired, the magnet was tuned, shimmed, and locked for D_2O in a fraction at the beginning of the HPLC fraction set that did not show a highly concentrated chromatographic signal. Even in fractions that have overlap in the mass detected HPLC signal, we can see separate signals in the NMR. One dimensional (1-D) proton (^1H) data was acquired using the noesypr1D pulse sequence (**Figure 2**). There were 4 dummy scans, and 256 scans collected with over 64,000 data points, acquired at 298K. The spectra were set to be centered on water and peak phasing was performed manually using Bruker TopSpin 3.6.4 software. A few fractions were reanalyzed, but minimal differences were seen in the new spectra, so the original NMR spectra was used.

The NMR data was transferred from Bruker TopSpin 3.6.4 NMR software to the Network of Advanced NMR (NAN) as it was being produced. The spectrometer collects data in a time domain known as the free induction decay (FID). NMRPipe, an NMR processing software, transformed

the FID to a Fourier Transform (FT), which showed this conversion in the frequency domain spectra (Wu et al., 2024). NMRPipe was used to reference the input spectra from the spectrometer to DSS at 0.0 ppm. The frequency domain consists of a series of peaks displayed by chemical shift values in parts per million (or ppm) along the x-axis and a level of relative intensity or power along the y-axis. Batch processing simplifies the processing of a large data set by automatically aligning and phasing the spectra as NMRPipe would if it was done manually. After the completion of the batch processing, the spectra are viewed, compared, and stacked in MestreNova Lab (MNova) (**Error! Reference source not found.Figure 2**). After viewing the spectra in MNova, there were not many changes to be made that would improve the spectra from the automatic batch processing.

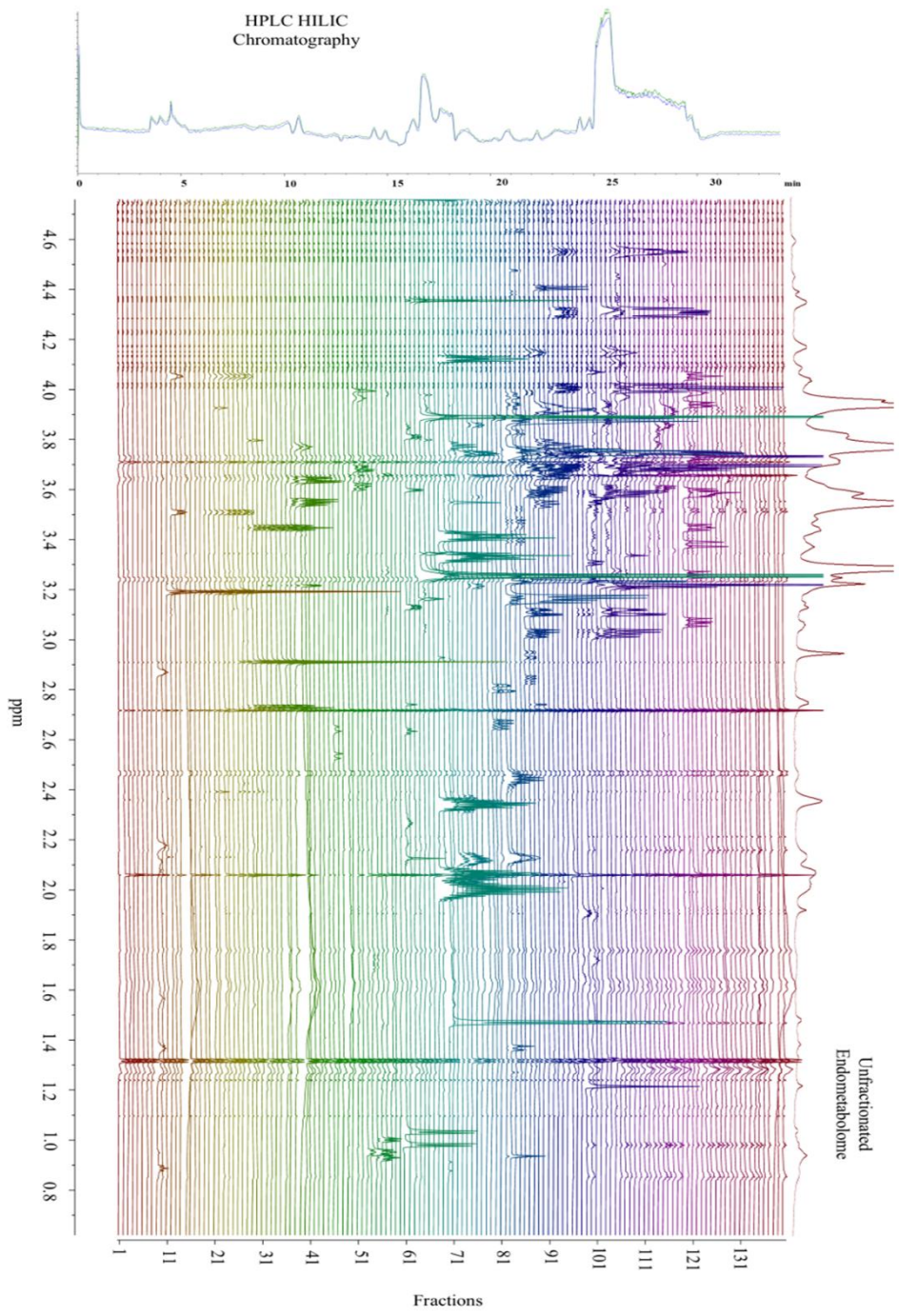

Figure 2 Stacked NMR Spectra of the T. pseudonana Endometabolome in 140 Fractions. *T. pseudonana* endometabolome separated into 140 fractions. **Far left:** HPLC HILIC Chromatography mass detected by the total ion count (TIC) at the given time. The chromatography also shows two identical chromatograms from two separate injections in which the 140 fractions were taken from. **Top:** Unfractionated *T. pseudonana* endometabolome ^{1}H-1D NMR spectra. **Middle:** 140 fractions of the separated *T. pseudonana* endometabolome. Each line and color represent a fraction. The variation in color shows differentiation between spectra from other spectra near it.

CHAPTER 3

RESULTS

3.1 Fraction Library

After the completion of the batch processing, the spectra were viewed, compared, and stacked in MestreNova Lab (MNova). The spectra showed a wide range of compounds in the endometabolome, which was expected since the endometabolome is very rich (Moran et al., 2016). In the HPLC chromatography, there are trends in intensity of compounds that gradually increase to near maximum detection and gradually decrease over consecutive fractions. The overlapping region at the end of the fractionation (about 24 minutes to 29 minutes) contained a concentrated number of metabolites, which can be seen in the stacked spectra from about Fraction 100 to Fraction 123. The unfractionated endometabolome chromatograph (Figure 2, top spectra) cannot show the level of detail that the fractionated spectra can. This evidence drives encouragement to use a fraction library for complex samples.

3.2 Annotation (SAFER)

The identification of compounds using NMR spectroscopy is not a straight path, as compounds may overlap in chromatography, fractionation and in NMR spectra (Edison et al., 2021). Researchers now have access to a great deal of virtual databases that provide chemical structures, NMR spectra (both real and simulated) and biological data for compounds (Walters, 2019). The Spectral Annotation by Feature Extraction and Reference (SAFER) matching approach [created by a former member of the Edison Lab, Michael T. Judge, Imperial College – London] allows a

set of NMR spectra to be matched with NMR reference spectra over metabolite databases containing more than 1,300 compounds (Judge & Martin, 2023). This approach uses the peak's shape, peak's definition (or intensity), and chemical shift ranges of associated peaks to match the acquired data to the compound referenced data. The SAFER approach can be useful but just like any other analysis, it produced limited results. There were positive matches, resulting in the high confidence of an identified compound but there were also some very low confidence matches and false matches that would require further analyses to validate. The SAFER reference compounds were derived using 500 MHz and 600 MHz magnets. Since the NMR fraction set was acquired using an 800 MHz magnet, simulated GISSMO (Guided Ideographic Spin System Model Optimization) reference spectra were used to compare compounds at 800 MHz as well. Potential SAFER matches were stacked with their GISSMO reference spectra. The initial attempt to identify compounds consisted of compiling a list of compounds found within *T. pseudonana* cell endometabolome in the scientific literature (Moran et al., 2022; Uchimiya et al., 2022), then using this list to initiate metabolite annotation (**Table 2**). Compounds in the list were cross checked with the compounds available in the SAFER compound pool. The SAFER matches were compared to the processed fraction data collected and the results were included in a table (**Table 3**). Figures 3 – 6 show matched compounds produced by SAFER. **Figure 3** (Proline) and **Figure 5** (L-Alanine) include compounds that were matches to the direct GISSMO database. **Table 3** does not include these two compounds as they were not matches detected by SAFER, but this could be a result of low detection within SAFER of compounds. For those compounds not included in the SAFER matches, I referred to compounds in **Table 2** and compared GISSMO simulated data to the processed data of the real fraction library using Mnova.

Table 2 Identified Endometabolite Compound List.

This is a *T. pseudonana* endometabolome compound list combined from identified metabolites in the scientific literature cited. This list is in alphabetical order.

Compounds from Scientific Literature		
-1,3-Glucan	Ethanolamine	Phosphoglycerate
3-Hydroxybutyrate	Folate	Phosphorylcholine
4-Hydroxybenzoate	Gluconate	Proline
4-Hydroxyphenylacetate	Glucose	Propionate
5'-Methylthioadenosine	Glucose-6-phosphate	Pyridoxal
7-Dehydrocholesterol	Glucosylglycerol	Pyridoxal phosphate
Acetate	Glutamate	Pyridoxine
Aconitate	Glutamine	Riboflavin
Adenosine	Glutathione	Ribose 5-phosphate
Adenosine monophosphate	Glycerol-3-phosphate	S-5'-adenosyl-L-homocysteine
Alanine	Glycerophosphocholine	S-adenosyl methionine
Aminobutyrate	Glycine	Sarcosine
Arachidonate	Glycine Betaine	Serine
Arginine	Gonyol	Spermidine
Asparagine	Guanine	Sucrose
Aspartate	Guanosine monophosphate	Sulfoactate
Beta(1,3)-glucan	Guanosine	Taurine
Biotin	Histidine	Thiamin
Caffeine	Homarine	Thiamin monophosphate
Chitobiose	Hydroxocobalamin	Threonine
Choline	Indole-3-acetamide	Thymidine
Citrate	Indole-3-acetate	Thymine
Citrulline	Inosine	Trehalose
Creatine	Isethionate	Trigonelline
Cyclic Guanosine Monophosphate	Isoleucine	Trimethylamine N-oxide (TMAO)
Cystathionine	Kynureine	Trimethylamine-N-oxide
Cysteate	Leucine	Tryptophan
Cysteinolate	Lysine	Tyrosine
Cytidine	Methionine	Uracil
Cytosine	Methyl indole-3-corboxylate	Uridine
Dihydroxypropane-sulfonate (DHPS)	N-(3-oxotetradecanoyl-L-homoserine lactone	Uridine diphosphate-glucosamine
Dimethylglycine	N-acetyl-galactosamine	Uridine diphosphate-glucose
Dimethylsulfonioacetate	Niacin	Valine
Dimethylsulfonio-propionate (DMSP)	Ornithine	Vanillate
Dimethylsulfoxonium propionate	Pantothenate	Xanthine
Ectoine	Phenylalanine	Xanthosine

Table 3 SAFER Matches by Visual Comparison. The below metabolite matches were produced by SAFER, using NMR compound reference libraries and ranked based on confidence in a visual comparison. The range in fractions in which SAFER identified the compound present is indicated above. The parts per million (or ppm) value and reference to processed spectral match confidence is indicated above as well.

Metabolite	Lower_Frac	Upper_Frac	Region (ppm)		Score Range			
L-lysine	98	102						
L-asparagine	85	88	2.791	4.033	0.22	0.20	*86	*87
Choline	11	13	3.17	4.102	0.24	0.23		
Choline	16	29	3.17	4.102	0.24	0.22		
Biotin	106	120	1.256	1.915				
Biotin	123	140	1.256	1.915				
Citrate	45	48	2.495	2.668	0.30	0.26	*46	*47
Citrate	53	55	2.495	2.668	0.29	0.27		
Dihydrouracil	31	39	2.593	3.517	0.74	0.67		
Dihydrouracil	26	30	2.623	3.488	0.41	0.33		
L-glutamic-acid	68	76	1.932	3.815	0.24	0.22		
Glutaric-acid	106	120	1.656	2.283	0.65	0.57		
Glutaric-acid	123	140	1.656	2.283	0.65	0.54		
Glutaric-acid	33	63	1.656	2.283	0.64	0.45		
3-4-Dihydroxy-L-phenylalanine	86	91	2.887	6.966	0.26	0.24		
3-4-Dihydroxy-L-phenylalanine	96	104	2.887	6.966	0.26	0.21		
D-3-phosphoglyceric-acid	87	90	3.761	4.265	0.55	0.63		
D-3-phosphoglyceric-acid	100	100	3.761	4.265	0.52	0.52		
D-3-phosphoglyceric-acid	102	102	3.761	4.265	0.65	0.65		
D-3-phosphoglyceric-acid	104	105	3.761	4.265	0.69	0.55		
Glycerol	37	39	3.503	3.868	0.58	0.57		
Glycerol	128	128	3.47	3.837	0.56	0.56		
L-threitol	86	88	3.529	3.754	0.37	0.34		
L-glutamine	71	76	2.05	3.81	0.48	0.43		
Butyric-acid	123	140	0.854	2.183	0.19	0.17		
Butyric-acid	106	114	0.854	2.183	0.19	0.18		
Butyric-acid	117	118	0.854	2.183	0.19	0.19		
Myo-inositol	119	120	3.214	4.084	0.29	0.25		
L-canavanine	68	70	2.142	4.157	0.28	0.27		

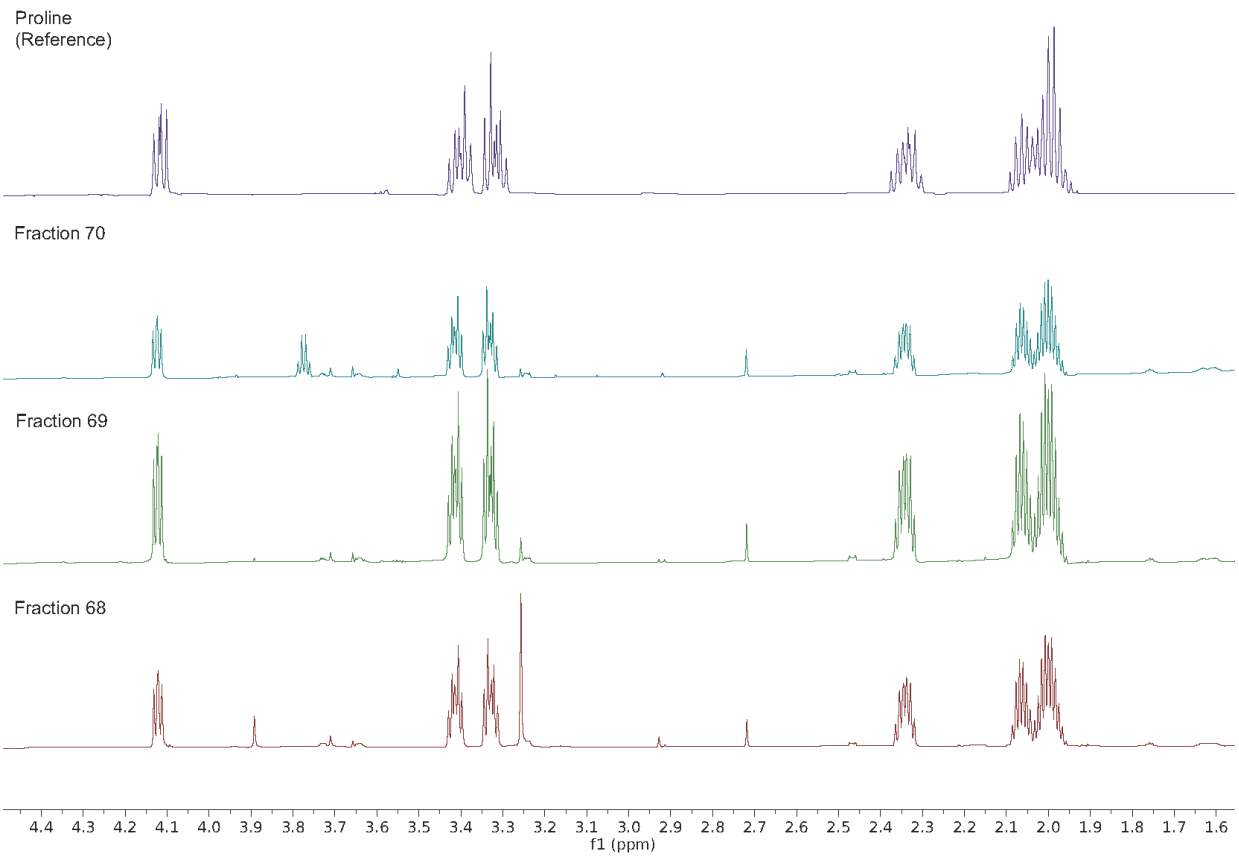

Figure 3 Proline SAFER Match. The reference compound Proline from reference libraries is compared to fractions from the fractioned *T. pseudonana* endometabolome set that show similar chemical shift values and peak shapes. The fractions contain other compounds, and this explains the additional peaks in the spectra. Fractions were collected using 100 mM Phosphate Buffer + 1/9 mM DSS in D_2O as the NMR solvent. Spectra are shown using an 800 MHz NMR magnet. The variation of color indicates the differentiations in fraction numbers.

SAFER suggested a match of the compound Proline being present in Fractions 68 - 70 (**Figure 3**). This serves as an example of a positive match from SAFER. It is visually evident that the referenced Proline from GISSMO is identical to that of the fractions shown, based on peak shape, intensity, and chemical shift values.

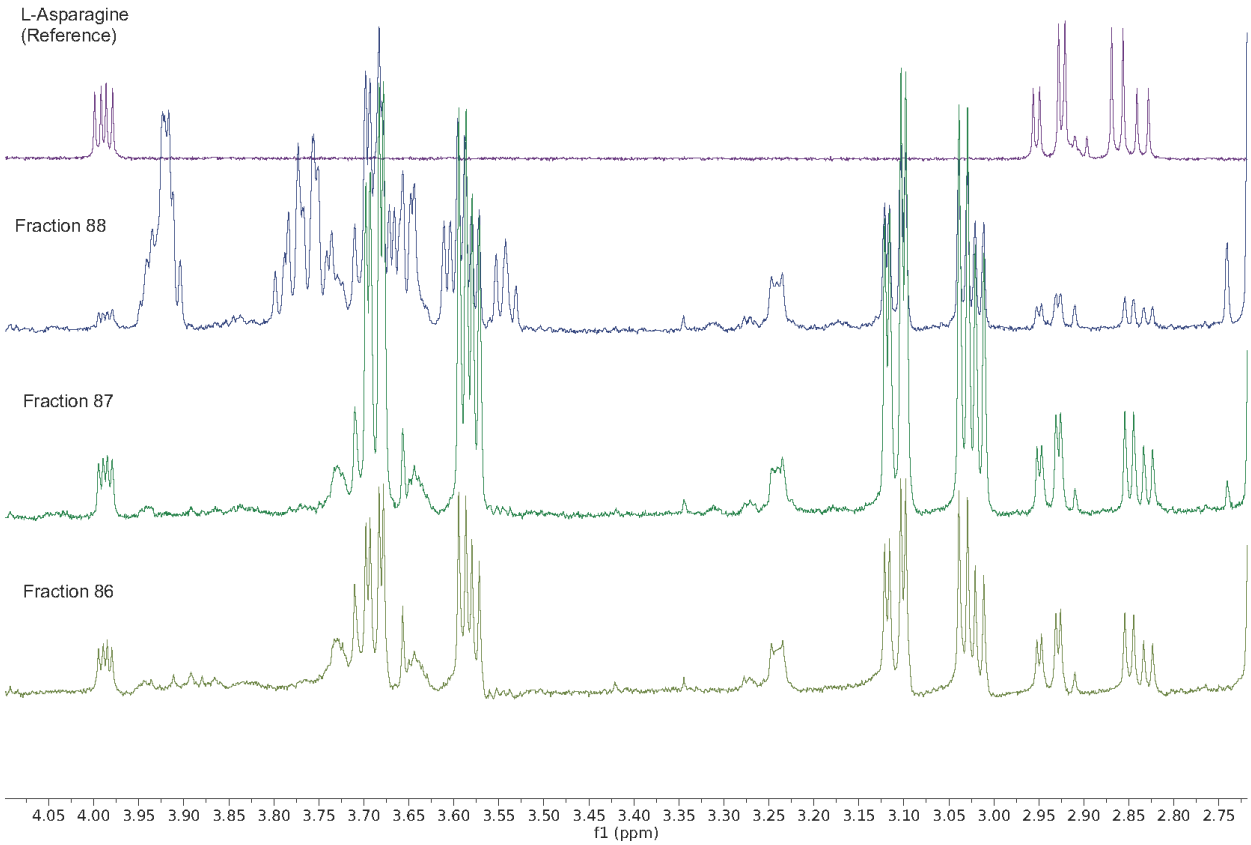

Figure 4 L-Asparagine SAFER Match. Fractions from the *T. pseudonana* endometabolome are compared to L-Asparagine from reference libraries show similar chemical shift values and peak shapes. The fractions contain other compounds, and this explains the additional peaks in the spectra. Fractions were acquired using 100 mM Phosphate Buffer + 1/9 mM DSS in D_2O as the NMR solvent. Spectra are shown using an 800 MHz NMR magnet. The variation of color indicates the differentiations in fraction numbers.

Fractions 86 - 88 showed a high confidence match of L-Asparagine (**Figure 4**) using SAFER. L-Asparagine is an example of showing peaks that may have been missed in annotation if the endometabolome was not fractionated.

Figure **2 Stacked NMR Spectra of the T. pseudonana Endometabolome in 140 Fractions.**

T. pseudonana endometabolome separated into 140 fractions. **Far left:** HPLC HILIC

Chromatography mass detected by the total ion count (TIC) at the given time. The chromatography also shows two identical chromatograms from two separate injections in which the 140 fractions were taken from. **Top:** Unfractionated T. pseudonana endometabolome ^{1}H-1D NMR spectra. **Middle:** 140 fractions of the separated T. pseudonana endometabolome. Each line and color represent a fraction. The variation in color shows differentiation between spectra from other spectra near it.

shows an abundance of peaks around the same chemical shift value (ppm) as L-Asparagine (about 3.99 ppm, 2.94 ppm, and 2.84 ppm). **Figure 1** and **Figure 2** also show that at about 24.00 minutes there is overlap in chromatographic signal. It becomes a challenge to receive this level of detail without defined separation of a complex sample.

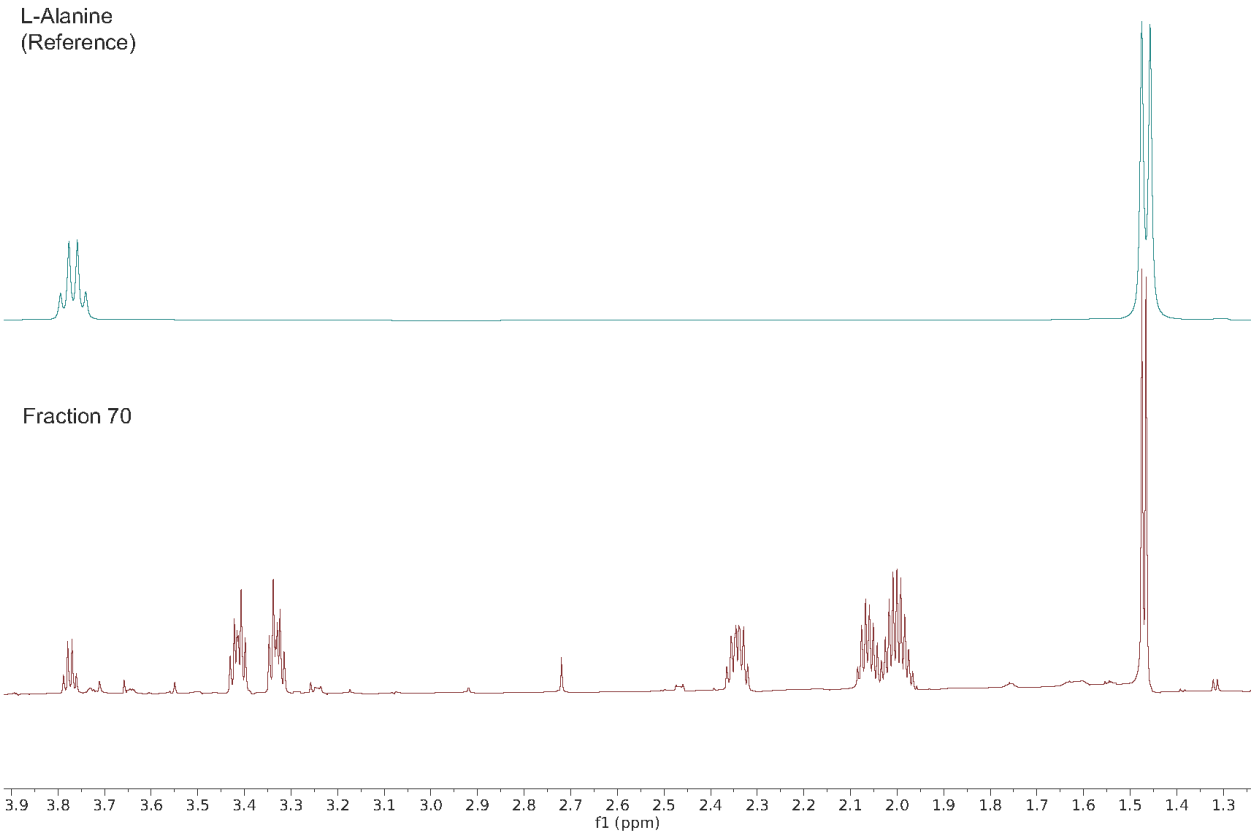

Figure 5 L-Alanine SAFE Match. Fractions from the *T. pseudonana* endometabolome are compared to L-Alanine from reference libraries show similar chemical shift values and peak shapes. The fractions contain other compounds, and this explains the additional peaks in the spectra. Fractions were collected using 100 mM Phosphate Buffer + 1/9 mM DSS in D_2O as the NMR solvent. Spectra are shown using an 800 MHz NMR magnet. The variation of color indicates the differentiations in fraction numbers.

L-Alanine is another example of a positive SAFER match (**Figure 5**). The fractions surrounding Fraction 70 showed little to no signal for the chemical shift of L-Alanine. This evidence supports that chromatographic fractionation has the potential to separate compounds based on polarity and we can suppose most of the L-Alanine compound in the endometabolome eluted into Fraction 70.

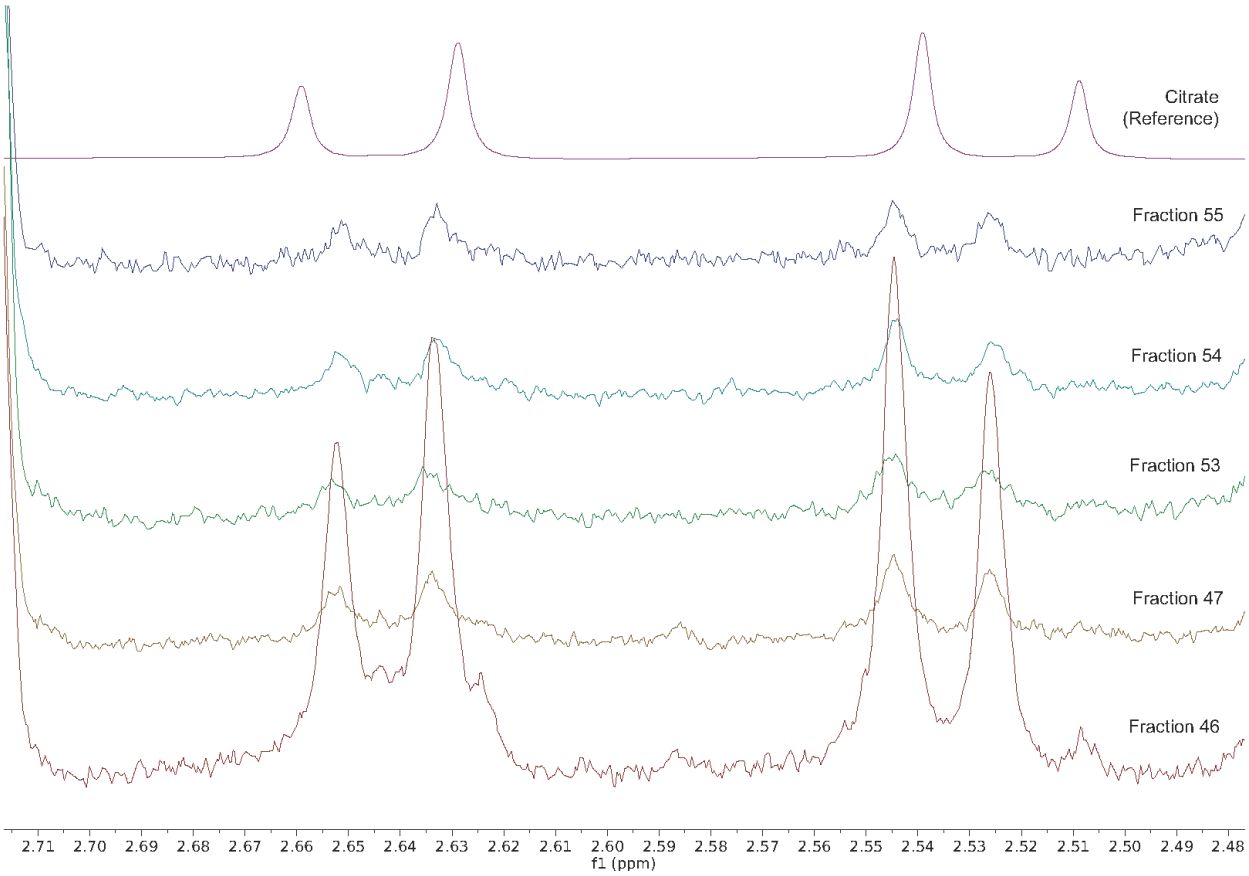

Figure 6 Citrate SAFER Match. Fractions from the *T. pseudonana* endometabolome are compared to Citrate from reference libraries show similar chemical shift values and peak shapes. The additional peaks in the figure are noise and have no relevance to the figure. Fractions were collected using 100 mM Phosphate Buffer + 1/9 mM DSS in D_2O as the NMR solvent. Spectra are shown using an 800 MHz NMR magnet. The variation of color indicates the differentiations in fraction numbers.

In other cases, SAFER matched compounds that are of lesser confidence. In **Figure 6**, the chemical shifts are within range (from the referenced SAFER spectra) but are not directly aligned compared to the other three compounds (Figures 4, 5, and 6). Differences in pH between fractions are believed to be the reason for this misalignment but further analysis would be required to validate. Fraction 46 shows what some may think is a positive correlation to the GISSMO

referenced Citrate, but the coupling constants are not identical. SAFER gave matches in Fractions 47, 53, 54, 55 but are questionable since their intensities are so low and could be considered noise (**Figure 6**). These fractions could not be labeled as positive matches without further analysis that are specific to Citrate and its composition.

It is important to reiterate the significance of limiting overlap when trying to identify compounds within a sample. Annotation would be limited and chaotic without chromatographic separation. Compounds can be overlooked or suppressed when they are not viewed to this level of detail. SAFER will remain a useful approach post NMR spectroscopy for direct annotation of metabolites. Performing multiple analyses and gathering supporting data from scientific literature is always encouraged to confirm matches generated by SAFER. SAFER matches allow researchers to start piecing together the puzzle of annotation in metabolomics, but it is still premature and dependent on adequate NMR features.

CHAPTER 4

FUTURE DIRECTIONS

The extraction of the endometabolome and the polar separation of its compounds provide a strong foundation for our understanding of phytoplankton metabolomics. With the fractionation of the endometabolome, experiments can be more specific for a fraction or range of fractions. Assays may require the presence of a specific metabolite and compound annotation across a small range of fractions also creates conditions where additional experiments can be conducted on the specified material. The annotation of metabolites in the exometabolome would be essential to this foundation but is a challenge to accomplish due to the low concentration of metabolites and saltiness of the sea water. Phytoplankton metabolites are consumed just as fast as they are released by heterotrophic bacteria. Both problems are difficult to control. There are various assays and functional analyses that can be utilized for understanding phytoplankton metabolomics, especially since the endometabolome has a higher concentration of metabolites. Deciding which experiments will be beneficial to your assay will depend on your target or goals (Wilson et al., 2020). Eash assay and analyses has variability, strengths, and limitations that will impact performance efficiency.

4.1.1 Endometabolome and Exometabolome Derivatization Comparison

There have been efforts to compare the endometabolome to the exometabolome of *T. pseudonana*. It is difficult to track the metabolites once they are released from the phytoplankton cells. The exometabolome metabolite extraction is different from that of the endometabolome. The

exometabolome must be derivatized due to the saltiness and instability of its compounds. The derivatization will potentially change the compound's composition, making it more difficult to complete any solid database matching. The hypothesis is that there are metabolites that overlap in the endometabolome and the exometabolome, meaning they are likely to have been released from the endometabolome to be present in the exometabolome. To test this hypothesis, it has been suggested to complete the benzoyl chloride derivatization on the endometabolome, conduct 1D-^{1}H NMR experiments, and then compare the derivatized NMR data to the normal, un-derivatized endometabolome data. This comparison would show what the metabolites should look like after derivatization, or pre- and post-released into the salt water. Utilizing the metabolites that were able to be derivatized, a comparison analysis could be completed on the derivatized exometabolome and endometabolome. An attempt on these experimentation for this hypothesis was done but could not be completed due to complications with the derivatization of the exometabolome. Potentially, it is a great comparison to be pursued.

4.1.2 Bacterial Growth Assay

Understanding the vital role that phytoplankton play in the marine food web is dependent on the release of metabolites needed for other organisms to survive (Edison et al., 2021). *Rugeria pomeroyi* is a bacterial that has a high affinity for metabolites from *T. pseudonana* (Uchimiya et al., 2022). Changes in the metabolic relationship of phytoplankton and bacteria will alter the marine food web and its environment (Edison et al., 2021). Phytoplankton and bacteria may even compete for nutrients in some cases (Seymour et al., 2017). Assuming the fractions of the phytoplankton endometabolome are rich in vital nutrients for these bacteria, these fractions can be fed to *R. pomeroyi* bacteria cultures and assayed based on their growth with the metabolites in

each fraction. Each culture would experience the same growth conditions, having the optimal density tested periodically. This experiment could advance understanding on how a range of fractions or specific metabolites in a fraction are necessary for successful heterotrophic bacteria growth.

Figure 7 shows where three additional endometabolome HPLC fractionations were repeated with endometabolome material from the same -80°C freezer stock as used for this project. When chromatographic signals were compared to the previous HPLC fractionation in **Figure 1**, they were identical. Using these fractions in bacterial growth assays, more compound specific experiments can be conducted the fractions. This assay also allows for a more targeted approach when annotating a specific range of fractions containing certain metabolites that are essential to heterotrophic bacteria, instead of trying to identify compounds in all 140 fractions. This assay has not been performed but may potentially be an effective experiment.

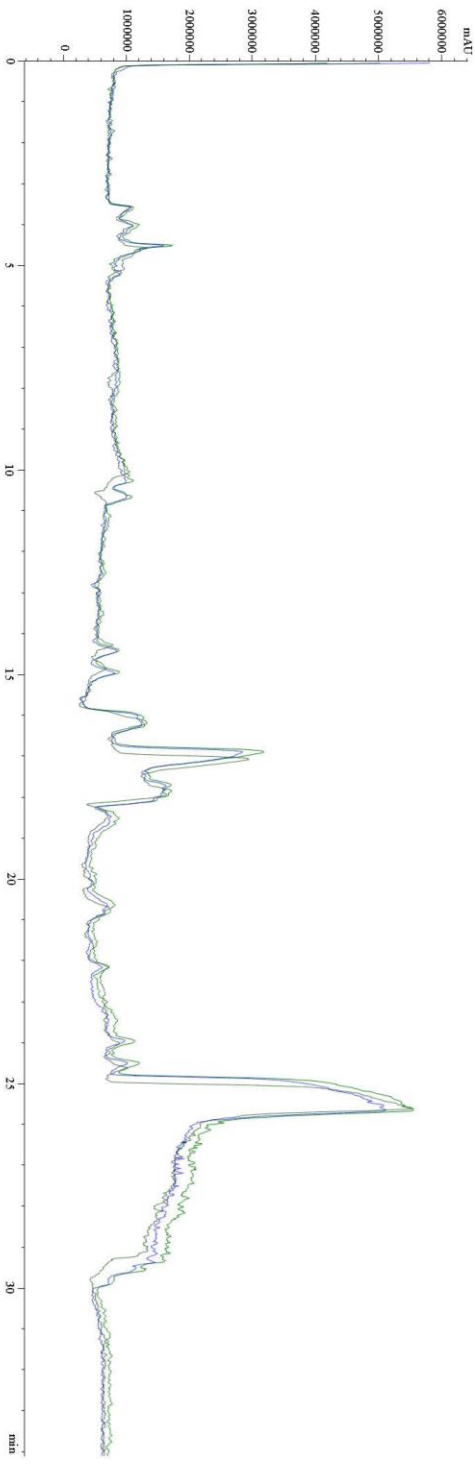

Figure 7 HPLC HILIC Chromatography for Bacteria Feeding Assay. The identical *T. pseudonana* endometabolome chromatograms of HPLC HILIC fractionation using the same mobile phases and gradient as shown in Table 1. 140 Fractions were collected. The chromatographic signals are mass detected by the total ion count (TIC) using a single quadrupole mass spectrometer. The injected sample is complex, resulting in signals reaching near maximum. The variation in color is to show the difference in chromatograms during identical injections.

4.2 Conclusion

As more contributions to marine science via phytoplankton and their metabolomics are made, new developments will strengthen endometabolome research. Creating a fraction library of metabolites can aid in the advancement and further understanding of metabolism in many different organisms. The extraction for *T. pseudonana* (CMP1335) endometabolome serves as the foundation for the fraction library and the fraction library will continue to expand. Ensuring all steps are followed in the extraction protocol, storage parameters, and mobile phases used, the method will remain highly through-put, supporting consistent sample preparation and HPLC fractionation. Efficient separation of compounds and consistency are essential to NMR metabolomic output for the fraction library. Using fractions in desired functional analyses to further investigate phytoplankton metabolomics can begin to answer detailed questions that would be difficult to answer without evidence. As there have been a few examples as to how a fraction library has been useful in metabolomics research, it is also important to understand how useful it would be in other scientific sectors of metabolomics. Whether the goal is to perform specific or diversified experiments, a fraction library is uniquely flexible and can fit into any project or research needs.

REFERENCES

Clement, R., Lignon, S., Mansuelle, P., Jensen, E., Pophillat, M., Lebrun, R., Denis, Y., Puppo, C., Maberly, S. C., & Gontero, B. (2017). Responses of the marine diatom Thalassiosira pseudonana to changes in CO(2) concentration: a proteomic approach. *Sci Rep*, *7*, 42333. https://doi.org/10.1038/srep42333

Durham, B. P., Sharma, S., Luo, H., Smith, C. B., Amin, S. A., Bender, S. J., Dearth, S. P., Van Mooy, B. A., Campagna, S. R., Kujawinski, E. B., Armbrust, E. V., & Moran, M. A. (2015). Cryptic carbon and sulfur cycling between surface ocean plankton. *Proc Natl Acad Sci U S A*, *112*(2), 453-457. https://doi.org/10.1073/pnas.1413137112

Edison, A. S., Colonna, M., Gouveia, G. J., Holderman, N. R., Judge, M. T., Shen, X., & Zhang, S. (2021). NMR: Unique Strengths That Enhance Modern Metabolomics Research. *Anal Chem*, *93*(1), 478-499. https://doi.org/10.1021/acs.analchem.0c04414

Falkowski, P. G., Barber, R. T., & Smetacek, V. (1998). Biogeochemical Controls and Feedbacks on Ocean Primary Production. *Science*, *281*(5374). https://doi.org/10.1126/science.281.5374.200

Ferrer-Gonzalez, F. X., Widner, B., Holderman, N. R., Glushka, J., Edison, A. S., Kujawinski, E. B., & Moran, M. A. (2021). Resource partitioning of phytoplankton metabolites that support bacterial heterotrophy. *ISME J*, *15*(3), 762-773. https://doi.org/10.1038/s41396-020-00811-y

Holderman, N. R., Ferrer-Gonzalez, F. X., Glushka, J., Moran, M. A., & Edison, A. S. (2023). Dissolved organic metabolite extraction from high-salt media. *NMR Biomed*, *36*(4), e4797. https://doi.org/10.1002/nbm.4797

Jonker, W., de Vries, K., Althuisius, N., van Iperen, D., Janssen, E., Ten Broek, R., Houtman, C., Zwart, N., Hamers, T., Lamoree, M. H., Ooms, B., Hidding, J., Somsen, G. W., & Kool, J. (2019). Compound Identification Using Liquid Chromatography and High-Resolution Noncontact Fraction Collection with a Solenoid Valve. *SLAS Technol*, *24*(6), 543-555. https://doi.org/10.1177/2472630319848768

Judge, M. T., & Martin, C. (2023). *EBI-Metabolights/SAFER-NMR: V2.1.2 (v.2.1.2.)*. https://doi.org10.5281/zenodo.10022483

Judge, M. T., Wu, Y., Tayyari, F., Hattori, A., Glushka, J., Ito, T., Arnold, J., & Edison, A. S. (2019). Continuous in vivo Metabolism by NMR. *Front Mol Biosci*, *6*, 26. https://doi.org/10.3389/fmolb.2019.00026

Moran, M. A., Kujawinski, E. B., Schroer, W. F., Amin, S. A., Bates, N. R., Bertrand, E. M., Braakman, R., Brown, C. T., Covert, M. W., Doney, S. C., Dyhrman, S. T., Edison, A. S., Eren, A. M., Levine, N. M., Li, L., Ross, A. C., Saito, M. A., Santoro, A. E., Segre, D.,…Vardi, A. (2022). Microbial metabolites in the marine carbon cycle. *Nat Microbiol*, *7*(4), 508-523. https://doi.org/10.1038/s41564-022-01090-3

Moran, M. A., Kujawinski, E. B., Stubbins, A., Fatland, R., Aluwihare, L. I., Buchan, A., Crump, B. C., Dorrestein, P. C., Dyhrman, S. T., Hess, N. J., Howe, B., Longnecker, K., Medeiros, P. M., Niggemann, J., Obernosterer, I., Repeta, D. J., & Waldbauer, J. R. (2016). Deciphering ocean carbon in a changing world. *Proc Natl Acad Sci U S A*, *113*(12), 3143-3151. https://doi.org/10.1073/pnas.1514645113

Olofsson, M., Ferrer-Gonzalez, F. X., Uchimiya, M., Schreier, J. E., Holderman, N. R., Smith, C. B., Edison, A. S., & Moran, M. A. (2022). Growth-stage-related shifts in diatom endometabolome composition set the stage for bacterial heterotrophy. *ISME Commun*, *2*(1), 28. https://doi.org/10.1038/s43705-022-00116-5

Seymour, J. R., Amin, S. A., Raina, J. B., & Stocker, R. (2017). Zooming in on the phycosphere: the ecological interface for phytoplankton-bacteria relationships. *Nat Microbiol*, *2*, 17065. https://doi.org/10.1038/nmicrobiol.2017.65

Siegel, D. A., DeVries, T., Cetinic, I., & Bisson, K. M. (2023). Quantifying the Ocean's Biological Pump and Its Carbon Cycle Impacts on Global Scales. *Ann Rev Mar Sci*, *15*, 329-356. https://doi.org/10.1146/annurev-marine-040722-115226

Uchimiya, M., Schroer, W., Olofsson, M., Edison, A. S., & Moran, M. A. (2022). Diel investments in metabolite production and consumption in a model microbial system. *ISME J*, *16*(5), 1306-1317. https://doi.org/10.1038/s41396-021-01172-w

Walters, W. P. (2019). Virtual Chemical Libraries. *J Med Chem*, *62*(3), 1116-1124. https://doi.org/10.1021/acs.jmedchem.8b01048

Widner, B., Kido Soule, M. C., Ferrer-Gonzalez, F. X., Moran, M. A., & Kujawinski, E. B. (2021). Quantification of Amine- and Alcohol-Containing Metabolites in Saline Samples Using Pre-extraction Benzoyl Chloride Derivatization and Ultrahigh Performance Liquid Chromatography Tandem Mass Spectrometry (UHPLC MS/MS). *Anal Chem*, *93*(11), 4809-4817. https://doi.org/10.1021/acs.analchem.0c03769

Wilson, B. A. P., Thornburg, C. C., Henrich, C. J., Grkovic, T., & O'Keefe, B. R. (2020). Creating and screening natural product libraries. *Nat Prod Rep*, *37*(7), 893-918. https://doi.org/10.1039/c9np00068b

Wishart, D. S., Cheng, L. L., Copie, V., Edison, A. S., Eghbalnia, H. R., Hoch, J. C., Gouveia, G. J., Pathmasiri, W., Powers, R., Schock, T. B., Sumner, L. W., & Uchimiya, M. (2022). NMR and Metabolomics-A Roadmap for the Future. *Metabolites*, *12*(8). https://doi.org/10.3390/metabo12080678

Wu, Y., Sanati, O., Uchimiya, M., Krishnamurthy, K., Wedell, J., Hoch, J. C., Edison, A. S., & Delaglio, F. (2024). SAND: Automated Time-Domain Modeling of NMR Spectra Applied to Metabolite Quantification. *Anal Chem*, *96*(5), 1843-1851. https://doi.org/10.1021/acs.analchem.3c03078

Zhang, L., Thevis, M., Piper, T., Jochmann, M. A., Wolbert, J. B., Kujawinski, D. M., Wiese, S., Teutenberg, T., & Schmidt, T. C. (2014). Carbon isotope ratio analysis of steroids by high-temperature liquid chromatography-isotope ratio mass spectrometry. *Anal Chem*, *86*(5), 2297-2302. https://doi.org/10.1021/ac403353x