Thermodynamic Modeling of Lipid Adaptation in Thermophilic Microbes

by

Grayson Maxwell Boyer

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2018 by the
Graduate Supervisory Committee:

Everett Shock, Chair
Hilairy Hartnett
Pierre Herckes

ARIZONA STATE UNIVERSITY

May 2018

ABSTRACT

Lipids perform functions essential to life and have a variety of structures that are influenced by the organisms and environments that produced them. Lipids tend to resist degradation after cell death, leading to their widespread use as biomarkers in geobiology, though their interpretation is often tricky. Many lipid structures are shared among organisms and function in many geochemical conditions and extremes. I argue it is useful to interpret lipid distributions as a balance of functional necessity and energy cost. This work utilizes a quantitative thermodynamic framework for interpreting energetically driven adaptation in lipids.

Yellowstone National Park is a prime location to study biological adaptations to a wide range of temperatures and geochemical conditions. Lipids were extracted and quantified from thermophilic microbial communities sampled along the temperature (29-91°C) and chemical gradients of four alkaline Yellowstone hot springs. I observed that decreased alkyl chain carbon content, increased degree of unsaturation, and a shift from ether to ester linkage caused a downstream increase in the average oxidation state of carbon (Z_C). I hypothesized these adaptations were selected because they represent cost-effective solutions to providing thermostable membranes.

This hypothesis was explored by assessing the relative energetic favorability of autotrophic reactions to form alkyl chains from known concentrations of dissolved inorganic species at elevated temperatures. I found that the oxidation-reduction potential (Eh) predicted to favor formation of sample-representative alkyl chains had a strong positive correlation with Eh calculated from hot spring water chemistry ($R^2 = 0.72$ for the O_2/H_2O redox couple). A separate thermodynamic analysis of bacteriohopanepolyol lipids found that predicted equilibrium abundances of observed polar headgroup distributions were also highly correlated with Eh of the surrounding water ($R^2 = 0.84$). These results represent the first quantitative thermodynamic assessment

of microbial lipid adaptation in natural systems and suggest that observed lipid distributions represent energetically cost-effective assemblages along temperature and chemical gradients.

This dissertation is dedicated to my wife Angela.

Thank you for your endless love and support.

ACKNOWLEDGEMENTS

I am deeply grateful for the guidance I received from my research advisor, Professor Everett Shock. He introduced me to the wonderful and mysterious biological and geochemical diversity of Yellowstone and challenged me to be the creative driving force of my own research. His enthusiasm, encouragement, and scientific insight shaped my transition from student to scientist. I am proud to be a member of his research group.

My advisory committee members Professors Hilairy Hartnett and Pierre Herckes have my deepest gratitude for their thoughtful feedback, hard-hitting questions, and attention to detail.

I extend many thanks to Professor Roger Summons for graciously allowing me to analyze lipids with the instrumentation in his laboratory at the Massachusetts Institute of Technology. I am indebted to Emily Matys for her assistance while running calibration curves and to Dr. Xiaolei Liu for kindly re-running a batch of lipid samples for me after I returned to Arizona. Dr. Florence Schubotz, now a researcher at the University of Bremen, was instrumental in providing analytical instruction, preparing authentic standards, and offering me helpful advice. Her work on lipid biogeochemistry continues to be a source of great inspiration. You have my thanks.

Dr. Alysia Cox and Dr. Jeff Dick, former postdocs of the Group Exploring Organic Processes In Geochemistry (GEOPIG), have my thanks as well. Alysia provided academic guidance that helped shape the organization of content in this dissertation. Jeff introduced me to coding in the R programming language and the R package he developed, CHNOSZ. His enthusiasm for integrating features into CHNOSZ and providing detailed answers to theoretical and technical questions proved extremely helpful.

There are a great many folks who contributed to the collection, processing, and analysis of the Yellowstone samples described in this dissertation. Most of these contributors are fellow members, alumni, or collaborators of GEOPIG at Arizona State University. First and foremost, I thank Jade Woods for her tireless assistance with lipid extractions over the span of six semesters. The work ethic, dependability, and professionalism that she displayed as a volunteer undergraduate research assistant made it a pleasure to work with her. Jenna Donatelli performed temperature, pH, and conductivity measurements, kept cartography records of the hot springs we visited, and lifted spirits with her upbeat attitude. Apar Prasad, Carlos Castellanos, and Kristin Johnson performed field spectrophotometry for water samples at Yellowstone National Park. As one of the few members of GEOPIG with an interest in thermodynamic calculations involving biomolecules, Apar has been a kindred spirit, and as such, his feedback has been valuable. Carlos is a great friend, and I'll never forget jamming out with him to music in the van on our 2011 and 2012 Yellowstone trips. Kristin and I joined GEOPIG at the same time, shared an interview with Everett Shock during our graduate school visitation, and ever since, Kristin has been a close friend. I have fond memories of exploring Tokyo, camping in Greer, Arizona, and trying out new board games with Kristin. Kris Fecteau managed, coordinated, and troubleshot ion chromatography, or in other words, did most of the heavy-lifting for the analysis of major dissolved ions. He also stomped me at chess in Yellowstone. Randall "Vince" Debes and Kirt Robinson had a great many responsibilities in the field and in the laboratory, including (but not limited to), collecting water and dissolved gas and preparing samples for analysis of major ions and dissolved organic/inorganic carbon (DIC/DOC). In addition to the countless roles they had in the successful analysis of water samples, Vince and Kirt have been inspirational friends and colleagues. I value the years of adventures we have had inside and outside

of work. Alta Howells, Michelle Santana, and Brian St. Clair deserve many thanks for analyzing dissolved gas and preparing samples for ion chromatography. I thank Tucker Ely and James Leong for their help with aqueous geochemical speciation; between the three of us, we created an automated tool with a custom thermodynamic database that greatly reduced the time it took to perform several crucial calculations. I also thank Panjai Prapaipong for leading the charge on trace metal analyses, and Natalya Zolotova for her guidance on DIC/DOC analyses. I am grateful for having the chance to work with you all.

I am deeply appreciative of my parents, Clayton and Lisa Boyer, for the love and encouragement they have shown me throughout my life. I consider myself fortunate to have been raised by parents from scientific and medical backgrounds. The knowledge they shared fostered my curiosity about science and the natural world. I also thank my parents-in-law, "Banana Joe" Joe and Cynthia Halasey, for their love and cheer through the years.

At the time of this writing, Angela Halasey and I have been together for fifteen years and married for six. Angela's unwavering love and support have shaped me in countless positive ways. I am incredibly fortunate to share my life with someone so wonderful.

TABLE OF CONTENTS

Page

LIST OF TABLES xi

LIST OF FIGURES xiii

CHAPTER

1 INTRODUCTION 1

2 THERMOPHILE LIPID OXIDATION STATE SUGGESTS BIOENERGETIC FAVORABILITY OF ALKYL CHAIN MODIFICATIONS ALONG TEMPERATURE AND REDOX GRADIENTS 14

2.1 Abstract 14

2.2 Introduction 15

2.3 Methods 19

2.3.1 Water chemistry 19

2.3.2 Sample collection and preparation 20

2.3.3 HPLC-MS 21

2.3.4 Mass spectral interpretation 22

2.3.5 Lipid quantification 23

2.3.6 IPL structural designations and chemical formulae 26

2.3.7 Calculation of average lipid properties and elemental composition 32

2.3.8 Calculation of IPL Z_C 33

2.3.9 Simulation of analytical uncertainty 34

2.4 Results 34

2.4.1 Water chemistry 34

2.4.2 Z_C of IPLs and their components 41

2.4.3 Abundance-weighted alkyl chain properties 48

CHAPTER Page

2.5 Discussion 55
2.6 Conclusion 60
3 THERMODYNAMIC FAVORABILITY OF THERMOPHILE LIPID CHAIN MODIFICATIONS ACROSS A TEMPERATURE AND REDOX GRADIENT 61
3.1 Introduction 61
3.2 Methods 68
3.2.1 Analysis of water chemistry 68
3.2.2 Analysis of environmental IPLs 69
3.2.3 Deriving properties and chemical formulae of $alkyl_{free}$ 69
3.2.4 Calculation of thermodynamic properties 72
3.2.5 Calculation of $alkyl_{free}$ relative abundance at hot spring conditions 82
3.2.6 Sensitivity analysis 86
3.3 Results 87
3.3.1 Water chemistry and Eh 87
3.4 Relative equilibrium abundances of $alkyl_{free}$ and predicted Eh_{alkyl} . . . 87
3.4.1 Sensitivity analysis of predicted equilibrium distributions . . . 94
3.5 Discussion 99
3.6 Conclusion 101
4 THERMODYNAMIC ASSESSMENT OF REDOX CONTROLS ON BACTERIOHOPANEPOLYOL HEADGROUPS 103
4.1 Introduction 103
4.2 Methods 109

CHAPTER Page

4.2.1 Water chemistry 109
4.2.2 Analysis of BHPs 110
4.2.3 Estimation of standard state thermodynamic properties of BHP polar headgroups. 111
4.2.4 Estimation of revised Helgeson-Kirkham-Flowers (HKF) equation of state parameters for BHP polar headgroups 118
4.2.5 Calculating BHP headgroup metastable equilibrium abundance and headgroup-predicted Eh 121
4.2.6 Calculating predicted fraction of aminopolyol BHP headgroups 121
4.3 Results 122
4.3.1 Observed BHP headgroup functionality 122
4.3.2 Z_C of BHP headgroups 124
4.3.3 Thermodynamic predictions of $Eh_{BHP,\ head}$ 127
4.3.4 Thermodynamic predictions of aminopolyol abundance 131
4.4 Discussion 135
4.5 Conclusion 139
5 CONCLUSION 141
5.1 Summary 141
5.2 Future Work 143
5.3 Broader Impacts 144
5.4 Concluding Remarks 146
REFERENCES 147
APPENDIX

CHAPTER Page

A Intact polar lipid response factors and mass spectral interpretation of novel structures 159

B Supplementary figures: predicted equilibrium abundances and sensitivity analyses of site-average free alkyl chains 174

C Supplementary figures: predicted equilibrium abundances of bacteriohopanepolyol polar headgroups 188

D Thermodynamic equations of state 196

LIST OF TABLES

Table Page

2.1 Observed polar lipids, assigned headgroup formulae, references used for identification, and assigned HPLC-MS quantification standards. 25

2.2 Selected geochemical and physical data from each sample site 35

2.3 Concentrations of selected redox-sensitive dissolved chemical species . . . 39

2.4 Weighted Z_C and chemical formula of IPLs and their component parts . 42

2.5 Summary of abundance-weighted alkyl chain properties 47

3.1 Summary of abundance-weighted structural characteristics of isoprenoidal alkyl chains. 70

3.2 Partial molal thermodynamic aqueous and hydration properties of aliphatic carboxylic acids, alcohols, and alkanes with nC $>$ 2 used in the estimation of saturated fatty acid, fatty alcohol, and fatty hydrocarbon properties. 74

3.3 Regression statistics for partial molal properties of fatty acids, fatty alcohols, and alkanes with nC $>$ 2 as a function of carbon length. 75

3.4 Partial molal thermodynamic aqueous and hydration properties of functional groups and chemical species used in the estimation of fatty amide properties. 77

3.5 Partial molal thermodynamic data for alkenes used in the estimation of unsaturated alkyl chain properties. 79

3.6 Partial molal thermodynamic data for second order groups used in the estimation of alkyl chain methyl branching, hydroxylation, and monolayer-spanning characteristics. 80

3.7 Partial molal thermodynamic data for pentacyclic and hexacyclic alkanes used in the estimation of GDGT internal ring properties. 81

Table Page

3.8 Estimated thermodynamic standard state partial molal properties and Helgeson-Kirkham-Flowers equation of state parameters for aqueous sample-averaged free alkyl chains. 83

3.9 Measured or calculated geochemical and physical data, along with Eh predicted from sample-averaged free alkyl chains. 88

4.1 Partial molal thermodynamic data for sugar alcohols used in the estimation of BHP polar headgroup properties 114

4.2 Partial molal thermodynamic properties of non-stereoisomeric sugar alcohols estimated from the linear regression of sugar alcohol diasteromer properties given in Table 4.1 with carbon length, as shown in Figure 4.3. 116

4.3 Partial molal thermodynamic data for second order groups used in the estimation of BHP headgroup properties. 117

4.4 Partial molal thermodynamic properties of aqueous chemical species used in the estimation of BHP aminopolyol headgroup properties. 119

4.5 Estimated thermodynamic standard state partial molal properties and Helgeson-Kirkham-Flowers equation of state parameters for aqueous BHP headgroups. 120

4.6 Select geochemical and physical data for samples, calculated and lipid-predicted redox potentials, and the average oxidation state of carbon in BHP headgroups. 125

A.1 HPLC-MS IPL standards and response factors 161

LIST OF FIGURES

Figure		Page
1.1	Cartoon cross-section of a lipid membrane bilayer	3
1.2	Five example chemical structures of bilayer-forming IPLs	4
1.3	Cartoon cross-section of a lipid membrane monolayer	5
1.4	Two examples of monolayer-forming intact polar lipids	7
1.5	Hypothetical net reactions to form a phospholipid and an aminolipid from a set of basis species	10
1.6	Hypothetical net reactions to form a lipid with a *cis*-double bond and a lipid with a cyclopropane ring	11
2.1	Structural designations used for IPL headgroups, backbones, and alkyl chains	30
2.2	Lipid alkyl chain modifications and backbone-chain linkage types organized by Z_C	38
2.3	Total concentrations of redox-sensitive aqueous chemical species in samples from Bison Pool, Mound Spring, Empress Pool, and Octopus Spring	40
2.4	Abundance-weighted Z_C of IPLs and their component parts	43
2.5	Relative abundance of backbone-alkyl chain linkage types	49
2.6	Number of aliphatic carbons (nC) per IPL alkyl chain as a function of temperature	51
2.7	Number of unsaturations (nUnsat) per IPL alkyl chain	52
2.8	Number of internal rings per GDGT	54
3.1	Hypothetical chemical structure of the sample-averaged free alkyl chain of Bison Pool sample BP1	64

Figure Page

3.2 Comparison of IPL backbone-alkyl chain linkage types and the free alkyl chains used in thermodynamic calculations 66
3.3 Predicted metastable equilibrium abundance of sample-averaged free alkyl chains of Bison Pool samples 91
3.4 Sensitivity analysis sample-averaged free alkyl chain metastable equilibrium abundance in Bison Pool samples 97
3.5 Alkyl chain-predicted Eh as a function of temperature, pH, and Eh 98
4.1 Generalized bacteriohopanepolyol chemical structure 106
4.2 Workflow used in the estimation of partial molal standard state thermodynamic properties of aqueous BHP headgroups 113
4.3 Linear regression of experimentally derived aqueous partial molal thermodynamic properties of sugar alcohols as a function of carbon length . 115
4.4 Relative abundances of BHP polar headgroup functionality observed in Yellowstone hot spring sediments and biofilms 123
4.5 Weighted Z_C of thermophile lipid components as a function of temperature and dissolved oxygen concentration 126
4.6 Predicted metastable equilibrium abundance of tetrafunctional and pentafunctional BHP headgroups in Bison Pool 130
4.7 Lipid-predicted Eh as a function of temperature and redox potential . . . 133
4.8 Predicted and observed fraction of aminopolyol BHP headgroups 134
4.9 Predicted fractional abundance of aqueous carbonate species as a function of pH 137
A.1 Mass spectra and putative structure for (N-methyl)glycosaminyl monoglycosyl phosphatidylarchaeol (MeNG-G-P-AR). 163

Figure Page

A.2 Mass spectra and putative structure for (N)glycosaminyl monoglycosyl phosphatidylarchaeol (NG-G-P-AR) 164
A.3 Mass spectrum and putative structure for (6-N,6-N,6-N)trimethyllysine lipid (TM-KL). 165
A.4 Mass spectra and putative structure for (6-N,6-N,6-N)trimethyllysine lipid (NG-GA-DAG). 166
A.5 Mass spectra and putative structure for unknown lipid '223'-DAG. 168
A.6 Mass spectrum and putative structure for glycosyl (N-methyl)glycosaminyl glycosyl phosphatidylarchaeol (G-MeNG-G-P-AR). 169
A.7 Mass spectra and putative structure for glycosyl (N)glycosaminyl glycosyl phosphatidylarchaeol (G-MeNG-G-P-AR) 170
A.8 Mass spectra and putative structure for triglycosyl (N-acetyl)glycosaminyl glycosyl dietherglycerol (3G-NAcG-G-DEG) 172
A.9 Mass spectra and putative structure for G-GA-DAG. 173
B.1 Metastable equilibrium abundance of sample-averaged free alkyl chains of Mound Spring 178
B.2 Metastable equilibrium abundance of sample-averaged free alkyl chains of Empress Pool 180
B.3 Metastable equilibrium abundance of sample-averaged free alkyl chains of Octopus Spring 181
B.4 Sensitivity analysis sample-averaged free alkyl chain metastable equilibrium abundance in Mound Spring samples 184
B.5 Sensitivity analysis sample-averaged free alkyl chain metastable equilibrium abundance in Empress Pool samples 186

Figure Page

B.6 Sensitivity analysis of sample-averaged free alkyl chain metastable equilibrium abundance in Octopus Spring samples 187
C.1 Predicted metastable equilibrium abundance of tetrafunctional and pentafunctional BHP headgroups in Mound Spring, Empress Pool, and Octopus Spring samples 195

Chapter 1

INTRODUCTION

Lipids, or fats, are biomolecules that perform many tasks essential to life. One such task is the formation of cell membranes, providing compartmentalization of metabolic pathways that would be unable to proceed otherwise, preserving disequilibrium in chemical potential required for harvesting energy, and separating the outside of the cell from the inside. Lipids host membrane-bound proteins, serve as anchors for cell walls, perform signaling, and store energy. Because lipids fill so many biological roles, and because organisms need these roles fulfilled in so many sets of environmental conditions, the variety of known lipid structures is immense (*e.g.,* Sturt *et al.*, 2004; Belin *et al.*, 2018; van Meer *et al.*, 2008; Yoshinaga *et al.*, 2011; Schouten *et al.*, 2013).

In this work, lipid function is mainly referenced in terms of the formation of lipid membranes to compartmentalize the cell or parts of the cell (*e.g.,* membrane-bound organelles in eukaryotes, thylakoids in cyanobacteria, the periplasmic space in Gram-negative bacteria, *etc.*). In all organisms, lipid membranes separate the inside of the cell from the outside. This barrier concentrates intracellular material to allow metabolic pathways to proceed, helps prevent harmful extracellular material from entering, and allows degrees of selectivity when importing or exporting material from the cell. Compartmentalization by lipid membranes also preserves the electrochemical gradients harnessed by organisms to harvest cellular energy, such as those produced by the electron transport chain. Compartmentalization is critically important for survival, and as such, biochemical pathways have emerged to regulate the stability of lipid membranes. The effectiveness of any given lipid membrane to compartmentalize adequately changes under different external conditions. For instance, decreasing

temperature often results in a decrease in the fluidity of lipid membranes, and *vice versa* for increasing temperature. If a lipid membrane becomes too rigid or too fluid, it may leak or rupture, disrupting the boundary between 'inside' and 'outside' and potentially causing great harm to the cell.

The physical properties of a lipid membrane depend on the chemical structures of the individual lipid molecules that comprise it. A simplified cartoon cross-section of a lipid membrane is depicted in Figure 1.1. The type of membrane depicted in this cartoon has two lipid layers, one at the top and one at the bottom, together referred to as a 'bilayer'. A typical lipid in the bilayer is also depicted schematically with its components labeled. These consist of a headgroup at one end, alkyl chains at the other end, and a backbone to connect them. The headgroup is typically polar and hydrophilic, or 'water-loving', while the alkyl chains are nonpolar and hydrophobic, or 'water-fearing'. The difference in polarity across the lipid and resulting non-covalent interaction with other polar lipid molecules is responsible for holding the lipid membrane together in an aqueous environment. Hydrophobic interactions tend to keep alkyl chains oriented toward the interior of the lipid membrane, while headgroups orient outward towards water on either side of the membrane. These lipids are sometimes referred to in the literature as intact polar lipids (IPLs). The word 'intact' is in reference to the goal of the analytical method, which is to analyze as much of the complete lipid as possible (in contrast to analyzing non-intact lipid components, such as hydrolyzed alkyl chains). Several examples of bilayer-forming IPLs with color-coded headgroups, backbones, and alkyl chains are depicted in Figure 1.2 to demonstrate that structural differences between lipids can occur in any component. For some organisms, such as many archaea, lipid membranes may be partially or exclusively composed of lipids that form a single layer, or 'monolayer' (depicted in Figure 1.3). Monolayer-forming IPLs also have polar headgroups, backbones, and

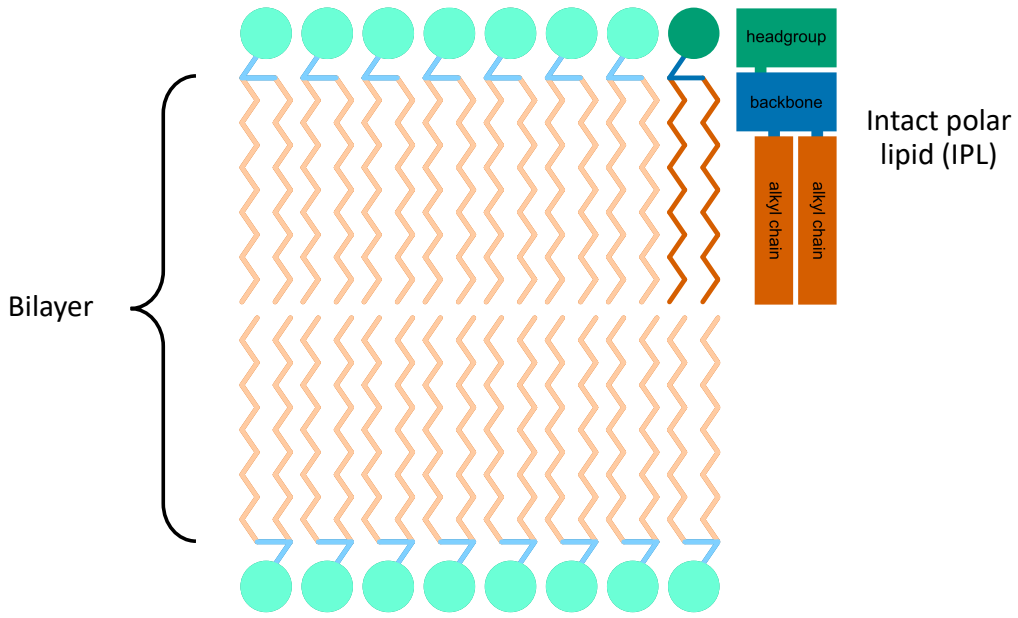


Figure 1.1: Cartoon cross-section of a lipid membrane bilayer. A single bilayer-forming intact polar lipid is highlighted. To the right of it is a simple schematic labeling common components of IPLs; headgroups (green), backbones (blue), and alkyl chains (orange).

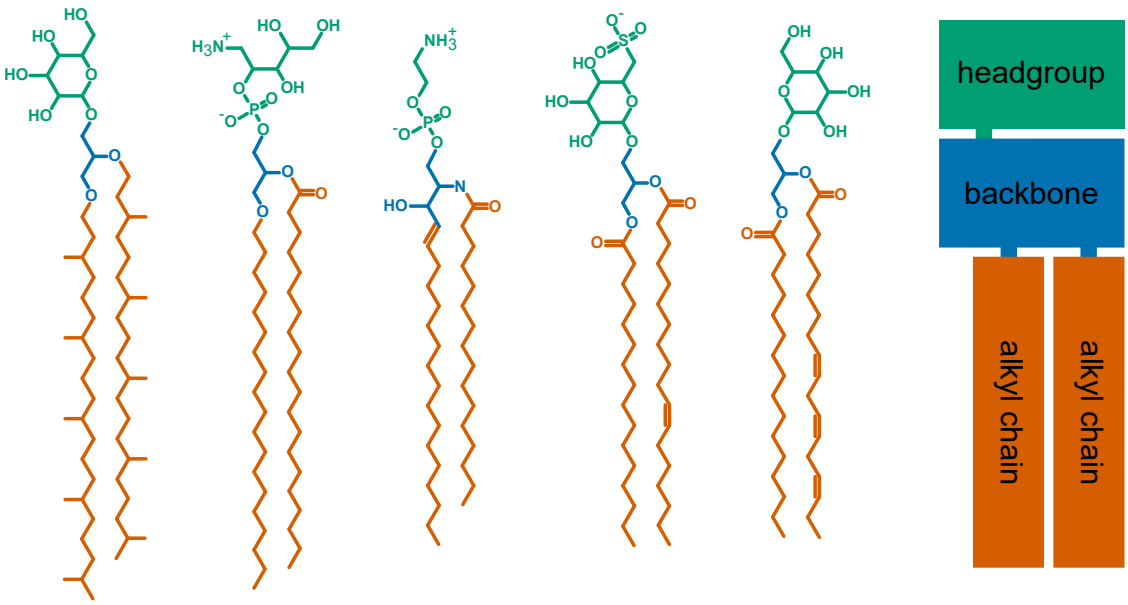


Figure 1.2: Five example chemical structures of bilayer-forming IPLs with a variety of headgroups (green), backbones (blue), and alkyl chains (orange).

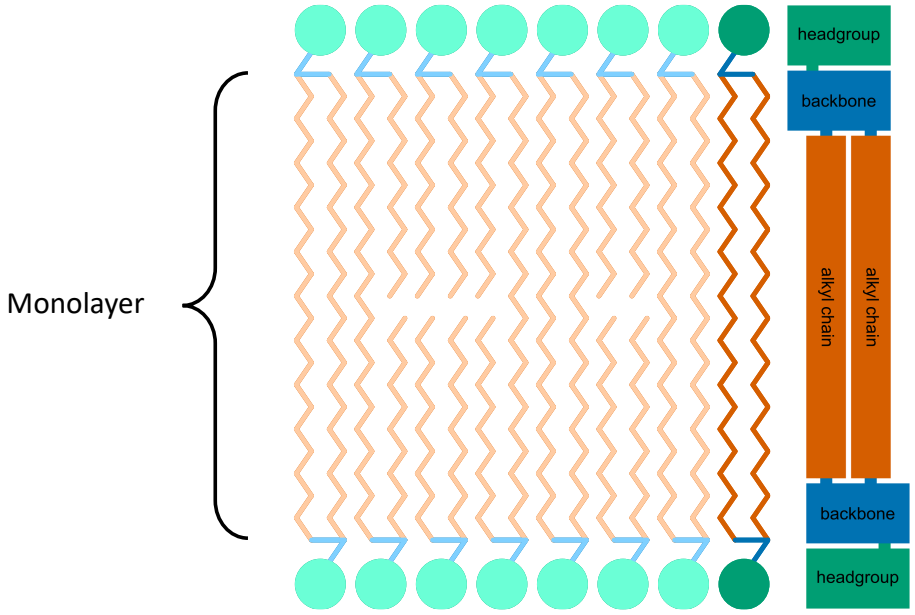


Figure 1.3: Cartoon cross-section of a lipid membrane that includes monolayer-forming lipids. A single monolayer-forming intact polar lipid is highlighted. To the right of it is a simple schematic labeling common components of monolayer-forming IPLs; headgroups (green), backbones (blue), and alkyl chains (orange).

nonpolar alkyl chains. Differences in structure between components can also exist for monolayer-forming lipids, as depicted by the examples in Figure 1.4.

A significant portion of the hydrophobic interior of cell membranes is composed of aliphatic hydrocarbons (*e.g.*, IPL alkyl chains). Many of these hydrocarbon structures are resistant to degradation and persist in the environment as recognizable pieces long after the death of the source organism. Some lipid structures are dated to billions of years old in the rock record (Brocks *et al.*, 2003a,b). Structural diversity, longevity, and potential traceability have led to the extensive use of lipids in biogeoscience as biomarkers for tracing sources of organic matter or predicting environmental conditions of the past. However, interpreting lipid biomarkers can be tricky because of the limited specificity of certain structures to a single organism or set of environmental variables, since many lipid structures are shared between organisms and function in a number of geochemical conditions, as discussed in Chapter 2.

I argue it is more useful to interpret patterns in lipid distributions in the context of the fitness advantage provided to their source organisms. For instance, an abundance of *cis*-double bonds in lipid chains might indicate that fitness is conferred by maintaining cell membrane fluidity in a low-temperature environment. However, the fitness advantage provided by a lipid structure, or any biomolecule for that matter, is not just determined by how well it functions, but also by its energetic cost. No matter how well a particular lipid structure might function, it must have an acceptable energetic cost to be an advantageous evolutionary investment. Furthermore, it seems likely that if several different evolutionary options are available for lipid structures that fill the same functional role with equal ability, the structure chosen by natural selection will be the one that saves the most cellular energy. Organisms synthesize lipids from the materials available to them, and it follows that an organism that makes the best use of available materials saves the most energy. Considering that

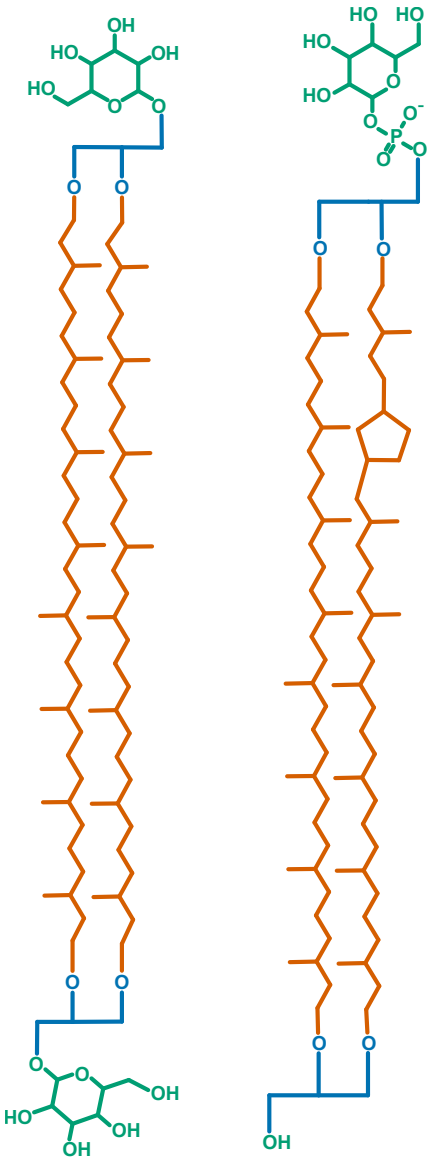


Figure 1.4: Two examples of monolayer-forming intact polar lipids. Lipid components are color-coded; headgroups are green, backbones are blue, and alkyl chains are orange.

functional lipids must be synthesized by organisms across the entire range of conditions inhabited by life and that these lipids must also have an acceptable biosynthetic cost, it is no wonder that so many lipid structures and compositions are observed.

The idea that lipid distributions are explained by natural cost-benefit analysis serves as the theoretical basis for this work. The major hypothesis of this work is that **microbes adapt their lipids to *maximize function* while also *minimizing cost***. By 'maximizing function', I refer to the process by which natural selection favors biomolecules, such as lipids, that provide a competitive advantage for their source organism by performing their biological tasks capably. This evolutionary process of maximizing lipid function was not investigated directly by this work, though it was assumed to be active within the sampled microbial communities. This assumption is based primarily on prior studies that tie lipid modifications to the regulation of membrane homeostasis. For instance, increasing lipid chain length and incorporating *cis*-double bonds to decrease or increase membrane fluidity and permeability, respectively (see review by Zhang and Rock, 2008). The incorporation of these various lipid modifications to produce stable, functional membranes can be interpreted as a maximization of lipid function. While most of this work frames lipid function in terms of regulating membrane homeostasis, the enhancement of other cell processes, such as the stabilization of photosystem proteins with phosphatidylglycerol and sulfoquinovosyldiacylglycerol lipids (Sato, 2004), can also be interpreted as lipid adaptation to maximize function.

Regardless of how well a lipid composition functions for an organism, it must have an acceptable biosynthetic cost. This work addresses the second half of the major hypothesis that relates to the 'minimization of cost'. By investigating the geochemical conditions influencing the potential cost of synthesizing lipids from available materials, I wanted to know whether patterns would emerge suggesting energetically driven

evolutionary convergence on observed lipid distributions. This led to my first sub-hypothesis, which is that **increasingly oxidized conditions will correspond to increasingly oxidized lipid distributions**, stemming from the idea that biosynthesis of oxidized biomolecules may be more cost-effective under oxidized conditions, and *vice versa* for reduced biomolecules under reduced conditions (Amend and Shock, 1998; Shock and Canovas, 2010; Dick and Shock, 2011, 2013). This sub-hypothesis is addressed by the work in Chapter 2, where I calculate the average oxidation state of carbon in IPLs and their components sampled along temperature and chemical gradients.

The second and third sub-hypotheses of this dissertation are variations of the major hypothesis but deal with individual lipid components. The second sub-hypothesis is that **adaptations to maximize the function of lipid alkyl chains also minimize cost** and the third sub-hypothesis is that **adaptations to maximize the function of lipid headgroups also minimize cost**. The decision to separately address the energetic costs of lipid alkyl chains and headgroups was made for reasons explained in greater detail in Chapters 3 and 4. The general approach used to estimate the relative energetic costs is based on modeling chemical reactions to form aqueous-phase lipid components from a chosen set of aqueous 'basis species' representing concentrations of inorganic solutes thought to be bioavailable to the primary producers of a microbial community, with the inclusion of electrons as a catch-all for the reducing power generated by the cell (*e.g.*, the reduced form of nicotinamide adenine dinucleotide phosphate, NADPH). Example reactions to form lipids from a set of basis species are shown in Figures 1.5 and 1.6. These represent net chemical reactions to form lipids from one possible set of basis species and intentionally do not take lipid synthesis pathways into account. Figure 1.5 depicts the formation of a phospholipid (top right) and an aminolipid (bottom right) from a set of basis

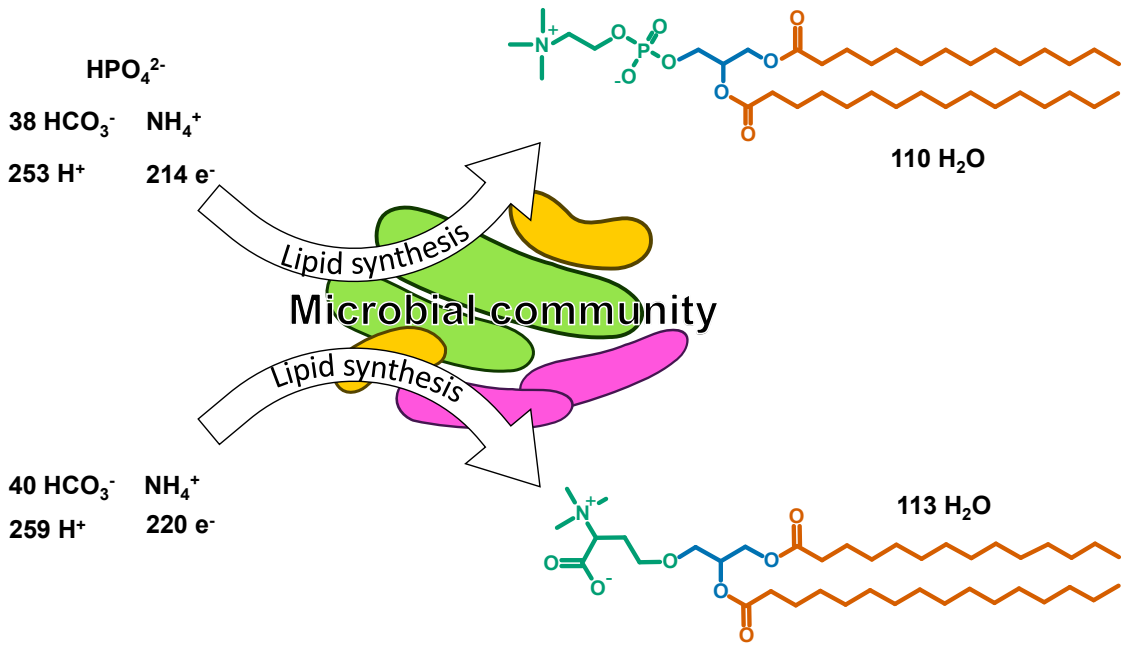


Figure 1.5: A hypothetical depiction of the net process of autotrophic uptake of nutrients and biosynthesis of a phospholipid (top right) and an aminolipid (bottom right) by a microbial community. Lipid headgroups, backbones, and alkyl chains are shown in green, blue, and orange, respectively, within the lipid chemical structures. The lipid structures depicted here differ only by their headgroups. In these reactions, lipids are formed from a set of basis species (left) that includes phosphate (HPO_4^{2-}), bicarbonate (HCO_3^-), ammonium (NH_4^+), protons (H^+), and electrons (e^-), with water (H_2O) as a byproduct.

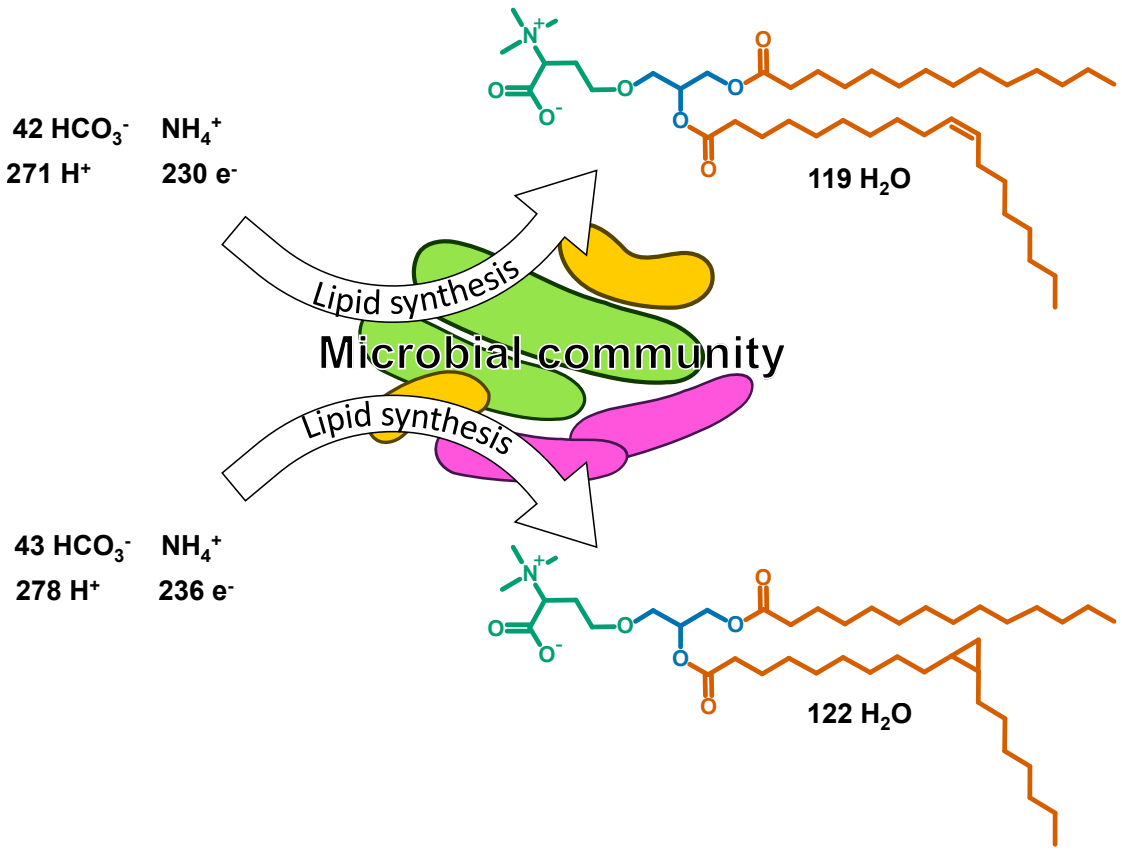


Figure 1.6: A hypothetical depiction of the net process of autotrophic uptake of nutrients and biosynthesis of a lipid with a *cis*-double bond in its alkyl chain (top right) and one with a cyclopropane ring (bottom right) by a microbial community. Lipid headgroups, backbones, and alkyl chains are shown in green, blue, and orange, respectively, within the lipid chemical structures. The lipid structures depicted here differ only by one alkyl chain. In these reactions, lipids are formed from a set of basis species (left) that includes bicarbonate (HCO_3^-), ammonium (NH_4^+), protons (H^+), and electrons (e^-), with water (H_2O) as a byproduct.

species (left) with water (right) as a product. Note that the stoichiometry of most reactants is different for both lipids, which only differ by headgroup structure, and that each reaction generates a different number of water molecules as a byproduct. If the chemical activities (which are related to concentration) and thermodynamic properties of the basis species and lipids have been measured or can be estimated, and the temperature of the system has been measured, then it becomes possible to calculate thermodynamic properties of these net reactions.

It is easy to imagine that net reactions to form the phospholipid and aminolipid in Figure 1.5 will have different thermodynamic properties depending on temperature or availability of basis species. Under phosphate replete conditions, the phospholipid might confer a greater fitness advantage than the aminolipid. However, if the availability of phosphate diminishes, the reaction to form the phospholipid might become so energetically unfavorable that its functional benefit is no longer worth the cost. Indeed, studies have demonstrated that under phosphorus starvation, some organisms can switch from producing phospholipids to aminolipids, including phytoplankton (van Mooy *et al.*, 2009; Martin *et al.*, 2011), bacteria (Minnikin *et al.*, 1974; Benning *et al.*, 1995) and fungi (Riekhof *et al.*, 2014). In the case of a community of microbes lacking the necessary genes to express alternate lipids during nutrient limitation, it is easy to imagine a scenario where a dip in phosphate availability would still result in a shift in lipid composition; microbes able to express lipids without phosphate would out-compete microbes reliant on phospholipids. The sub-hypothesis that adaptations to maximize the function of lipid headgroups also minimize cost is explored further in Chapter 4.

Next, consider the two lipid variants shown in Figure 1.6 and the net reactions to form them from basis species. These lipids differ in the way a bend is introduced into an alkyl chain, which is produced by the incorporation of a *cis*-double bond in

one lipid (top right) and a cyclopropane ring in the other (bottom right). There may be a functional advantage for an organism to incorporate a *cis*-double bond rather than a cyclopropane ring, but assume for now that each modification results in an equally capable lipid. Note that there are differences in the stoichiometry of the reactants and products in the net reactions to form the two lipid structures, and that the thermodynamic properties of the two lipids likely differ. Depending on the temperature and chemical conditions of the system, it is reasonable that one lipid structure is more cost-effective than the other, though guessing which is arguably not as straightforward as the example with phospholipids and phosphate starvation in Figure 1.5. There may not be a great difference in energy required to produce a lipid with a double bond versus a cyclopentane ring, though I would argue that even small energetic savings would add up significantly when considering how many individual lipids an organism needs for its membranes. The thermodynamic calculations described in Chapter 3 were performed to address how lipid alkyl chain modifications might maximize lipid function while minimizing energetic cost.

This work represents the first step in placing lipid adaptation into a quantitative thermodynamic context that allows prediction from temperature and chemical composition. While this dissertation deals with communities of thermophilic microbes, I hope the bioenergetic principles described herein will contribute to enhanced understanding of lipid distributions in a greater geochemical context.

Chapter 2

THERMOPHILE LIPID OXIDATION STATE SUGGESTS BIOENERGETIC FAVORABILITY OF ALKYL CHAIN MODIFICATIONS ALONG TEMPERATURE AND REDOX GRADIENTS

2.1 Abstract

Distributions of microbial lipids change across thermal and geochemical gradients, coinciding with changes in community composition and the drive to preserve membrane homeostasis and thermostability. In this study, I extracted lipids from microbial biomass collected along the thermal and redox gradients of four hot springs in Yellowstone National Park (YNP) and calculated the abundance-weighted average oxidation state of carbon, Z_C, of intact polar lipids (IPLs) and their component headgroups, backbones, alkyl chains. I found that carbon in full IPLs and their alkyl chains becomes more oxidized downstream of each of the four hot spring sources, coinciding with decreased water temperature and increased concentrations of oxidized inorganic solutes such as dissolved oxygen. IPLs are most reduced in the hot, reduced conditions upstream, with weighted Z_C between -1.68 and -1.56. This value increases gradually downstream to around -1.36 to -1.33 in microbial communities living between 29.0-38.1°C. This near-linear increase in abundance-weighted Z_C along outflow channels can be attributed to adaptations in alkyl chains to regulate membrane fluidity and permeability; namely, a shift from ether-linked to ester-linked alkyl chains, a decrease in the number of aliphatic carbons per chain (nC), an increase in the number of unsaturation per chain (nUnsat), and an increase in the number of internal rings in archaeal tetraether lipids. Each of these modifications increases the Z_C of alkyl

chains, and by extension, full IPLs. Weighted Z_C of IPL backbones showed no trends with temperature or redox state. The Z_C of IPL headgroups decreased downstream but had a relatively small effect on the Z_C of full IPLs. A Monte Carlo-style analysis was performed to test the sensitivity of observed Z_C trends. Over the course of 999 iterations, liquid-chromatography mass-spectrometry (LC-MS) integrated peak areas were varied randomly up to $\pm 30\%$ and response factors by two orders of magnitude. Observed correlations in abundance-weighted Z_C remained strong for full IPLs and alkyl chains, but not for headgroups, which were sensitive to random noise at downstream sites. The downstream increase in the weighted Z_C of thermophile IPLs and their alkyl chains suggests a bioenergetic evolutionary advantage to solving the problem of membrane thermostability with reduced alkyl chain modifications under reduced conditions and oxidized modifications under oxidized conditions.

2.2 Introduction

Organisms have adapted compositions of membrane lipid structures, as well as ways to modify these compositions and structures, to regulate fluidity, structural integrity, permeability, and overall functionality in response to temperature, pressure, and chemical composition of their surroundings. Nature has utilized a number of strategies to allow membrane function across the entire range of conditions in which life exists, such as altering lipid headgroup and backbone structures, incorporating sterols, hopanoids, alkenones, or membrane-spanning monolayer-forming lipids such as glycerol dialkyl glycerol tetraethers (GDGTs), or other molecules, modifying chain length, methyl branching, the number of unsaturations and hydroxylations, and using various combinations of chain-backbone ester, ether, and amide linkage types (Marlowe *et al.*, 1984; Belin *et al.*, 2018; van Meer *et al.*, 2008).

Attempts to tie a particular lipid trait to a specific organism or geochemi-

cal variable are often met with surprising exceptions and caveats. For example, 2-methylhopanoids were once thought to be exclusive biomarkers for cyanobacteria (Summons *et al.*, 1999). This assumption was reevaluated after the discovery of these lipids in an anoxygenic phototroph (Rashby *et al.*, 2007); since then, 2-methylhopanoids have been linked to a variety of bacterial phyla (Ricci *et al.*, 2014). The lipid crenarchaeol, once thought to be a biomarker for planktonic marine archaea and a proxy for marine temperature, was later found in Nevada hot springs between 40 and 84°C with no temperature correlation (Pearson *et al.*, 2004). Thermophilic archaea were the only known source of GDGT lipids until it was discovered that GDGTs were widespread in environments below 20°C (Schouten *et al.*, 2000). Furthermore, GDGTs have now been traced to non-archaea with the discovery that anaerobic bacteria in peat may produce non-isoprenoidal GDGTs (Weijers *et al.*, 2006). The number of rings in GDGT alkyl chains is frequently used in 'ring indices' as a proxy for temperature (Schouten *et al.*, 2002; Tierney, 2012), though conflicting reports have led to questions about the temperature-dependence of GDGT ring expression (Sollich *et al.*, 2017). As more is learned about lipid sources and expression, it is becoming increasingly clear that many structural adaptations in lipids function in a wide range of natural systems.

Take, for instance, the intriguing parallels in polar lipid distributions sampled spatially across two very different environments with extreme chemical gradients; 1) the relatively isothermal redox gradient of a Black Sea water column and sediment with depth (Schubotz *et al.*, 2009; Schröder, 2015), and 2) the wide temperature range spanned by terrestrial hot spring outflow channels (Schubotz *et al.*, 2013, 2015). Lipids found in the shallow oxic zone of the Black Sea tend to be shorter-length, relatively unsaturated, and bilayer-forming with primarily ester-linked alkyl chains; the same characteristics also describe lipids abundant in lower-temperature, downstream

areas of a hot spring outflow channel. In the higher-temperature, upstream areas of the hot spring, monolayer-forming GDGTs are abundant and lipid alkyl chains tend to be relatively saturated, ether-linked, and composed of a greater number of aliphatic carbons. These are also the characteristics of lipids abundant in the anoxic sediment of the Black Sea.

It follows that observed lipid structural adaptations provide functional membranes for both hot spring thermophiles and the mesophiles of the Black Sea and that these structures were selected because they confer a fitness advantage to their source organisms within their respective set of environmental conditions. It should be noted, however, that temperature is almost certainly not the only environmental stress governing distributions of lipid structural adaptations in thermophiles. As is often the case in hydrothermal systems, temperature gradients coincide with gradients in water chemistry, such as pH, salinity, solute concentrations, and oxidation-reduction (redox) potential.

My goal was to quantify the influence of temperature and water chemistry on lipid adaptation. Specifically, I wanted to know whether there was a correlation between environmental redox potential and the oxidation state of carbon in lipids forming thermostable membranes under these conditions, and hypothesized that **increasingly oxidized conditions correspond to increasingly oxidized lipid distributions**. To approach this, I extracted lipids from eighteen sediment and microbial biomass samples collected across the thermal and redox gradients of four YNP hot springs: Bison Pool (BP), Mound Spring (MS), Empress Pool (EP), and Octopus Spring (OS). These locations were selected to improve the probability that observed lipid distributions are the result of adaptation to temperature and redox chemistry by minimizing potential influence of salinity gradients. Lipids were chromatographically separated by high-performance liquid chromatography (HPLC) and analyzed

by accurate-mass tandem mass spectrometry (MS/MS), affording details about the structural characteristics of IPLs and their component headgroups, backbones, and alkyl chains. This analysis opens the possibility of examining how these parts and their modifications correlate with one another, or with the temperature, chemical composition, and redox state of the surrounding environment.

Chemical formulae of IPLs and their components permit the calculation of Z_C as a metric for lipid oxidation state, where Z_C is the oxidation state of individual carbons averaged across an entire molecule. A relatively oxidized molecule has a Z_C that is more positive than that of a reduced one. Lipids have some of the most reduced carbon compared to other biomolecules since a significant portion of their structure is composed of 'fatty' hydrocarbons responsible for forming the hydrophobic interior of the membrane (Likens, 2010). To the best of my knowledge, the work described in this dissertation marks the first time that Z_C has been calculated for full IPL structures and their component parts.

I found that along the thermal and redox gradients of all four hot springs, the abundance-weighted Z_C of thermophile IPLs increased gradually (more oxidized carbon) from upstream to downstream, *i.e.*, away from the hot spring source, coinciding with decreased temperatures and increased oxidation potential of dissolved inorganic solutes, such as O_2. Furthermore, I found that changes in alkyl chain modifications, such as the number of aliphatic carbons, the degree of unsaturation, chain-backbone linkage chemistry, and inclusion of rings in the alkyl chains of glycerol dialkyl glycerol tetraethers (GDGTs), were most responsible for the observed increase in lipid oxidation state downstream.

To test the sensitivity of the observed trends in abundance-weighted Z_C, and to partially simulate sources of analytical uncertainty, an additional analysis was performed where lipid abundances were varied randomly over the course of 999 iterations.

Regardless of this treatment, resulting distributions of weighted Z_C continued to suggest IPLs and their alkyl chains are becoming more oxidized downstream in all four hot spring outflow channels.

The gradual increase in the oxidation state of lipids along thermal and redox gradients of terrestrial hot springs may suggest that alkyl chains adaptations to maintain membrane fluidity and thermostability utilize chemical structures that are energetically cost-effective within the redox conditions experienced by their source microbial communities.

2.3 Methods

2.3.1 Water chemistry

Temperature, conductivity, and pH were measured in the field as close to sampling locations as possible and before sample collection. Temperature and conductivity were measured with a YSI model 30 meter. Sample pH was measured with a WTW brand 3300i or 3110 model pH meter with WTW probe calibrated daily with pH 4, 7, and 10 buffer solutions at ambient temperature. Concentrations of dissolved oxygen and total sulfide were obtained from unfiltered water samples in the field using a Hach 2400 or 2800 portable spectrophotometer with Hach reagents and protocols. Water samples collected for laboratory analyses were filtered in the field with Supor^{TM} (Pall Corporation) 0.2 μm polyethersulfone (PES) syringe filters into 30 mL HDPE Nalgene bottles and stored at -20°C. Concentrations of total ammonium, nitrate, nitrite, and sulfate were obtained by ion chromatography on two Dionex DX-600 systems; one for the analysis of cations and the other for anions. Suppressor columns on both systems were regenerated with deionized water to improve the signal-to-noise ratio. The anion analysis system was equipped with a potassium hydroxide eluent gener-

ator, carbonate removal device, and AS11-HC/AG11-HC columns. Columns were equilibrated with 5 mM hydroxide for 10 minutes before each injection. The injection volume was 100 μL for anions. Using a constant flow rate of 1.0 mL/min, the eluent hydroxide concentration was held isocratically at 5 mM for 5 minutes, then increased over the course of 31 minutes with a nonlinear gradient (Chromeleon curve 8). The cation analysis system was equipped with CS-16 and CG-16 columns. Cation samples were acidified with 6 N methanesulfonic acid (MSA) to 19 mM final concentration. The injection volume was 75 μL for cation analysis. The columns were eluted isocratically with 19 mM MSA and a flow rate of 0.5 mL/min. Ion concentrations were obtained by comparison to calibration curves created using mixed ion standards (Environmental Express, Charleston, SC, USA). Quantification accuracy was verified by the inclusion of mixed ion-quality control standards (Thermo Scientific, Waltham, MA, USA) before, between, and after samples in each tray.

2.3.2 Sample collection and preparation

Sediment and biofilms for lipid analysis were collected with sterilized spatulas or forceps into sterile specimen containers and frozen on dry ice in the field before storage in a -80°C freezer. Frozen samples were freeze-dried and homogenized with a sterile mortar and pestle. Lipid extractions were carried out using a modified version of the Bligh and Dyer procedure (White and Ringelberg, 1998). Briefly, 0.5-2 g sediment or 200-800 mg biofilm was dissolved in a mixture of methanol (MeOH), dichloromethane (DCM), and 50 mM phosphate buffer at pH 7.4 (2:1:0.8 v/v). The mixture was sonicated for 10 minutes and then centrifuged for 10 minutes at 2000 rpm. The supernatant was collected and the remaining sediment or biofilm underwent one more extraction with the same solvent proportions, followed by two more times with a mixture of MeOH, DCM, and 50 mM trichloroacetic acid buffer at pH 2

(2:1:0.8) to aid extraction of GDGTs (Nishihara and Koga, 1987), and one more time with a mixture of 3:1 DCM:MeOH to account for less polar lipids. A liquid-liquid extraction was performed by adding equal volumes of water and DCM to the pooled supernatant, which was then mixed and allowed to separate into aqueous polar and organic nonpolar phases. The nonpolar organic phase was collected and the remaining aqueous phase was washed with equal parts DCM for two additional rounds of lipid-liquid extraction. The resulting total lipid extract (TLE) was dried under N_2 and redissolved in 9:1 DCM and MeOH.

2.3.3 HPLC-MS

Aliquots of the TLE were chromatographically separated on an Agilent 1200 series high-performance liquid chromatograph (HPLC) equipped with a Waters Acquity Ultra Performance Liquid Chromatography ethylene bridge hybrid (BEH) amide column according to the hydrophilic interaction chromatography (HILIC) method described in Wörmer *et al.* (2013). Mobile phases included solvent A, a mixture of acetonitrile, DCM, formic acid, and ammonia (750:250:0.015:0.15 v/v) and solvent B, a mixture of MeOH, H_2O, formic acid, and ammonia (500:500:4:4). The initial eluent was 99% solvent A and 1% solvent B that was brought to 5% B with a linear gradient over 4 minutes. The gradient continued to 25% B over 18.5 minutes, then to 40% over 0.5 minutes and held isocratically for 3.5 minutes. The flow rate was held constant at 0.4 mL min^{-1} throughout each run. Mass spectral analysis of IPLs was performed in positive ion mode on an Agilent 6520 Accurate-Mass Quadrupole Time-of-Flight (Q-TOF) mass spectrometer with an electrospray ionization source.

2.3.4 Mass spectral interpretation

IPLs were identified by the exact mass (M) of their parent ion (*i.e.*, the parent lipid molecule with either a proton adduct, $[M + H]^+$, ammonium ion adduct, $[M + NH_4]^+$, or no adduct, $[M]^+$, in the case of IPLs bearing headgroups with inherently charged quaternary amines), and by comparing mass-to-charge (m/z) fragmentation patterns to previously published data as described in Sturt *et al.* (2004). Table 2.1 summarizes references used for structural elucidation or mass spectral interpretation of IPLs. Structures exceeding the analytical window ($m/z > 2000$) were not detected, potentially leading to the exclusion of higher molecular weight lipids.

While headgroup moiety identities, number of chains, backbone-chain linkage types, unsaturations, and aliphatic chain carbons were inferred based on mass spectra, other structural information was not obtained, such as carbon positions of double bonds or hydroxylations on alkyl chains or the positions of glycosidic bonds in sugar headgroup moieties. The presence and position of chain branching in isoprenoidally derived chains were inferred based on carbon number. Branching in non-isoprenoidal chains, such as those found in iso- and anteiso fatty acids, could not be determined from their straight-chain counterparts, necessitating 'number of aliphatic carbons per chain' as a metric of alkyl chain carbon content rather than 'chain length', which implies distance spanned by straight or branching chains. Non-GDGT alkyl chain cyclizations, such cyclopropane fatty acids synthesized by certain bacteria (Grogan and Cronan, 1997), could not be discerned from unsaturations, as both types of chain modification have two fewer hydrogen atoms relative to a saturated straight chain, and, as such, were counted as unsaturations. It is important to note that knowledge of the carbon positions of alkyl chain modifications, or whether a 2Da loss is from unsaturation or cyclization, and other fine details of molecular configuration are not

required for the calculation of lipid Z_C. Z_C depends solely on elemental abundances, oxidation states of non-carbon elements, and molecular charge.

2.3.5 Lipid quantification

IPLs were quantified based on manual peak integration of identified parent ions. The mole fraction of the i^{th} IPL in a sample, x_i, was calculated using

$$x_i = \frac{I_i \cdot RF_i^{-1} \cdot mi_i^{-1}}{\sum_i \left(I_i \cdot RF_i^{-1} \cdot mi_i^{-1} \right)}, \tag{2.1}$$

where I_i stands for the manually integrated MS peak area, RF_i indicates the analytical response factor chosen, and mw_i designates the monoisotopic mass, all taken for the i^{th} IPL parent ion.

Analytical response factors were applied in this study to partially account for differences in IPL ionization efficiency. Response factors were estimated by taking the linear slope of the injected masses versus peak intensity for a small suite of co-analyzed IPL standards purchased from Avanti Polar Lipids. Because authentic standards are not available for every observed IPL structure, response factors were assigned based primarily on the similarity of headgroups to those of existing standards, under the assumption that headgroups are the chemical feature most likely responsible for differences in ionization efficiency among observed IPLs. Authentic standards and response factors are shown in Table A.1 and their assignments to observed IPLs are reported in Table 2.1. IPLs with headgroup analogs among standards include monoglycose (1G), diglycose (2G), sulfoquinovose (SQ), phosphatidylinositol (PI), diphosphatidyl glycerol (DPG), phosphatidylcholine (PC), phosphatidyl (N,N-dimethyl)ethanolamine (PDME), phosphatidylethanolamine (PE), phosphatidylglycerol (PG), phosphatidyl (N-methyl)ethanolamine (PME), and phosphatidylserine (PS). The rationale guiding response factor assignments for IPLs with no direct headgroup analog are

briefly outlined below. IPLs with nitrogen-bearing groups, such as diglycosyl (N-acetyl)glycosaminyl glycose (2G-NAcG-G), triglycosyl (N-acetyl)glycosaminyl glycose (3G-NAcG-G, monoglycosyl (N)glycosamine (G-NG), (N)glycosaminyl glycoronic acid (NG-GA), glycosyl (N-methyl) glycosaminyl glycosyl phosphate (G-MeNG-G-P), glycosyl (N)glycosaminyl glycosyl phosphate (G-NG-G-P), (N-methyl)glycosaminyl glycosyl phosphate (MeNG-G-P), (N-acetyl)glycosaminyl phosphate (NAcG-P), (N)glycosaminyl monoglycosyl phosphate (NG-G-P), aminophosphopentanetetrol (APT), or unidentified '223' headgroups, as well as ceramide (CER) lipids were assigned the response factor obtained from C42:0 PC diacylglycerol (DAG), under the assumption that the charge assumed by a nitrogen-bearing group during mass spectral analysis may cause these lipids to have somewhat comparable ionization efficiencies. This assumption was based on qualitative assessment of the differences in relative peak intensities of nitrogen-bearing and non-nitrogen bearing standards. The former has peak intensities of about one order of magnitude higher than the latter, on average. Lipids with a glycoronic acid (GA) headgroup, a non-nitrogen-containing single-moiety glycosyl group, were assigned the response factor obtained from a 1G-DAG standard composed of a mixture of C34:2 and C34:3 chain lengths. Lipids with a monoglycosyl glycoronic acid (G-GA), triglycose (3G), or tetraglycose (4G) headgroup were assigned the response factor of the standard containing the closest number of non-nitrogen-bearing glycosyl moieties; a 2G-DAG standard mixture of C34:2, C34:3, and C36:6 chain lengths. Glycophospholipids with no nitrogen, such as monoglycosyl phosphate (1G-P), diglycosyl phosphate (2G-P) or triglycosyl phosphate (3G-P) headgroups, were assigned the response factor obtained from C32:0 PI-DAG. All GDGTs were assigned the response factor obtained from the 1G-GDGT-PG standard that included a mixture of H-shaped and non-H-shaped alkyl chains with 0-3 internal rings. All aminolipids, including ornithine lipids (OL), monohydroxylated

Table 2.1: Observed polar lipids, assigned headgroup formulae, references used for identification, and assigned HPLC-MS quantification standards.

Headgroup		Backbone-chain	Ref‡	RF§	Headgroup		Backbone-chain	Ref‡	RF§
Abbreviation*	Formula†	linkage types*			Abbreviation*	Formula†	linkage types*		
Glycolipids					*Phospholipids*				
1G	$C_6H_{11}O_6$	DEG, AEG, DAG	a, b	1	APT	$C_5H_{13}NO_7P$	DEG, AEG, DAG	a, b	3
		GDGT	b, c	2	DPG	$C_3H_8O_9P_2$	DAG×2	a	7
		CER	a, d	3	PC	$C_5H_{14}NO_4P^+$	DAG	a, b	3
2G	$C_{12}H_{21}O_{11}$	DEG, AEG, DAG	a, b	4	PDME	$C_4H_{11}NO_4P$	DAG	j	8
		GDGT	b, c	2	PE	$C_2H_7NO_4P$	DAG, DEG, CER	a, c, d	9
3G	$C_{18}H_{31}O_{16}$	DEG, DAG	a	4	PG	$C_3H_8O_6P$	DAG	a, b	10
		GDGT	b, c	2	PME	$C_3H_9NO_4P$	DAG	j	11
2G-NAcG-G	$C_{24}H_{41}NO_{20}$	DAG, DEG	b, e	3	PS	$C_3H_7NO_6P$	DAG	a	12
3G-NAcG-G	$C_{30}H_{51}NO_{25}$	DEG	f	3					
4G	$C_{24}H_{41}O_{21}$	GDGT	b, c	2	*Aminolipids*				
GA	$C_6H_9O_7$	DAG	g	1	BL	$C_7H_{15}NO_3^+$	DAG	b, k	13
G-GA	$C_{12}H_{19}O_{12}$	DAG	f	4	OL	$C_5H_{11}N_2O_2$	FAHFA	l, b	13
G-NG	$C_{12}H_{22}NO_{10}$	DAG, DEG	h	3	OL-1OH	$C_5H_{11}N_2O_2$	FAHFA-OH	g, l	13
NG-GA	$C_{12}H_{20}NO_{11}$	DAG, AEG, DEG	f	3	TM-KL	$C_9H_{20}N_2O_2^+$	FAHFA	f	13
SQ	$C_6H_{11}O_8S$	DAG	b	5	TM-OL	$C_8H_{18}N_2O_2^+$	FAHFA	m	13
					TM-OL-1OH	$C_8H_{18}N_2O_2^+$	FAHFA-OH	m	13
Glycophospholipids					*Unidentified*				
1G-P	$C_6H_{12}O_9P$	GDGT	b, c	2	'223'	$C_7H_{12}NO_7$	DAG	f	3
2G-P	$C_{12}H_{22}O_{14}P$	DEG	a, c	6					
		GDGT	b, c	2					
3G-P	$C_{18}H_{32}O_{19}P$	GDGT	b, c	2					
G-MeNG-G-P	$C_{19}H_{35}NO_{18}P$	DEG	f	3					
G-NG-G-P	$C_{18}H_{33}NO_{18}P$	DEG	f	3					
MeNG-G-P	$C_{13}H_{25}NO_{13}P$	DEG	f	3					
NAcG-P	$C_{12}H_{21}N_2O_{11}P$	DAG, DEG	b, i	3					
NG-G-P	$C_{12}H_{23}NO_{13}P$	DEG	f	3					
PI	$C_6H_{12}O_9P$	DAG, AEG, DEG	a, b	6					
		CER	a, b, d	3					

* see text for abbreviations

† formulas correspond to elemental abundances contained in green 'headgroup division' boxes of IPLs depicted in Figure 2.1

‡ references used for structural elucidation and/or mass spectral interpretation; a. Sturt *et al.* (2004); b. Schubotz *et al.* (2013) ; c. Yoshinaga *et al.* (2011); d. Karlsson *et al.* (1998); e. Ferreira *et al.* (1999); f. this work (see Appendix A); g. Diercks *et al.* (2015); h. Schubotz *et al.* (2015); i. Yang *et al.* (2006); j. Wang *et al.* (2015); k. Benning *et al.* (1995); l. Zhang *et al.* (2009); m. Moore *et al.* (2013)

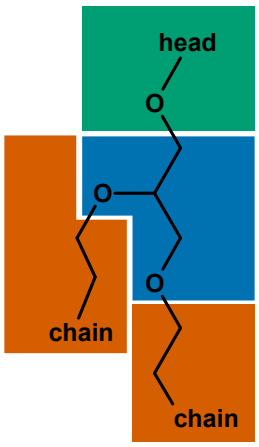
§ IPL standard used to determine analytical response factor for IPL quantification. All standards were purchased from Avanti Polar Lipids unless noted otherwise; 1. mix of C34:6 and C36:6 1G-DAG; 2. 1G-GDGT-PG; 3. C42:0 PC-DAG; 4. mix of C34:2, C34:3, C34:6, and C36:6 2G-DAG; 5. mix of C34:2, C34:3, and C36:6 SQ-DAG; 6. C32:0 PI-DAG; 7. C72:4 DPG; 8. C32:0 PDME-DAG; 9. C32:0 PE-DAG; 10. C32:0 PG-DAG; 11. C32:0 PME-DAG; 12. C32:0 PS-DAG; 13. C32:0 DGTS-d9

ornithine lipids (OL-OH), trimethylornithine lipids (TM-OL), monohydroxylated trimethylornithine lipids (TM-OL-OH), trimethyllysine lipids (TM-KL) and betaine lipids (BL), were assigned the response factor obtained from C32:0 1,2-dipalmitoyl-sn-glycero-3-O-4'-[N,N,N-trimethyl(d9)]-homoserine (DGTS-d9) based on structural similarities of their amino acid headgroups.

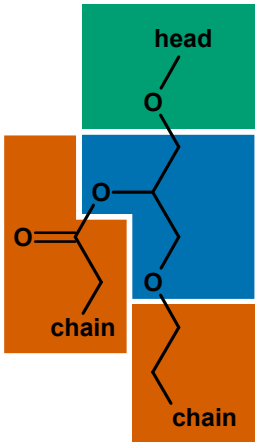
2.3.6 IPL structural designations and chemical formulae

Headgroups, backbones, and alkyl chains serve as the three basic building blocks comprising IPL structure, each with its own set of observed structural variations. The schematic used to categorize divisions among observed headgroup-backbone-alkyl chain variation is shown in Figure 2.1, with green, blue, and orange boxes indicating headgroup, backbone, and alkyl chain structures, respectively. When considering differences across widely varying lipid structures, it is important to define strict boundaries between components for the sake of consistently comparing properties that depend on chemical formulae, such as Z_C or the number of aliphatic carbons in an alkyl chain. This was particularly important for the comparison of IPLs with a 'typical' glycerol backbone to those without one, such as 1,2-alkanediols (Figure 2.1h), CER-lipids (Figure 2.1d), fatty acid esters of hydroxy fatty amides (FAHFAm, Figure 2.1e), and monohydroxylated FAHFAm-lipids (Figure 2.1f).

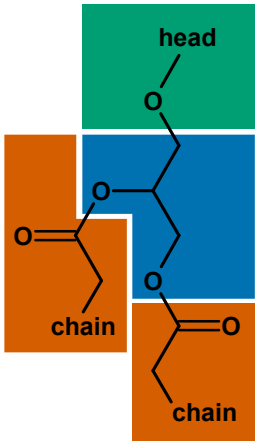
The portion of an IPL comprising the headgroup was structurally designated as one or more covalently bonded polar moieties linked to one or more backbones, represented by structures contained within green boxes in Figure 2.1. In some cases, the headgroup itself may be directly linked to one or more alkyl chains, such as in the tentative structures shown in Figures 2.1i and 2.1k. To calculate consistent chemical formulae, headgroups include electronegative atoms linking them to backbones or chains, such as the oxygen atom that forms the glycosidic bond between backbone and



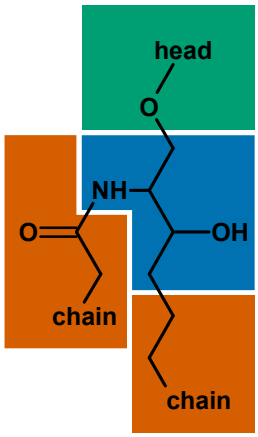
(a) DEG



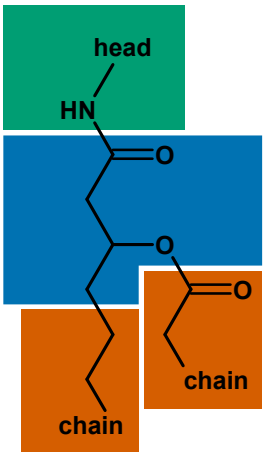
(b) AEG



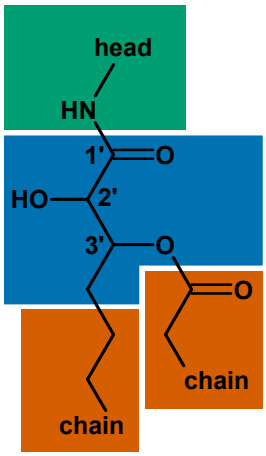
(c) DAG



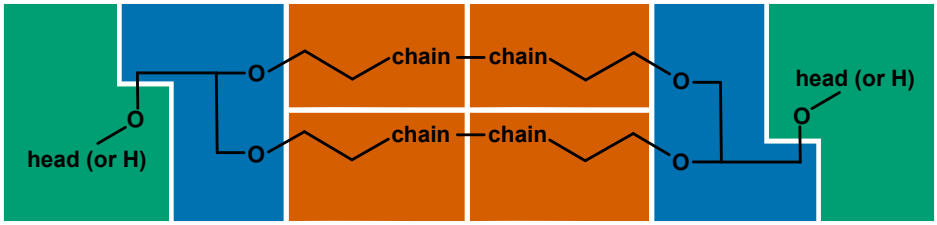
(d) CER



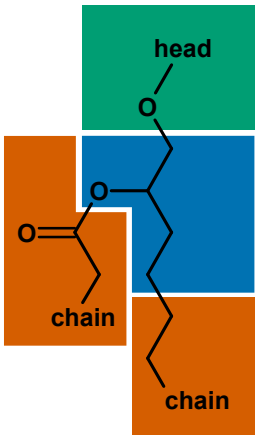
(e) FAHFAm



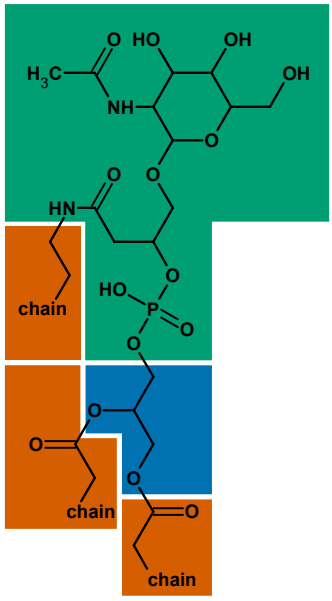
(f) FAHFAm-OH*



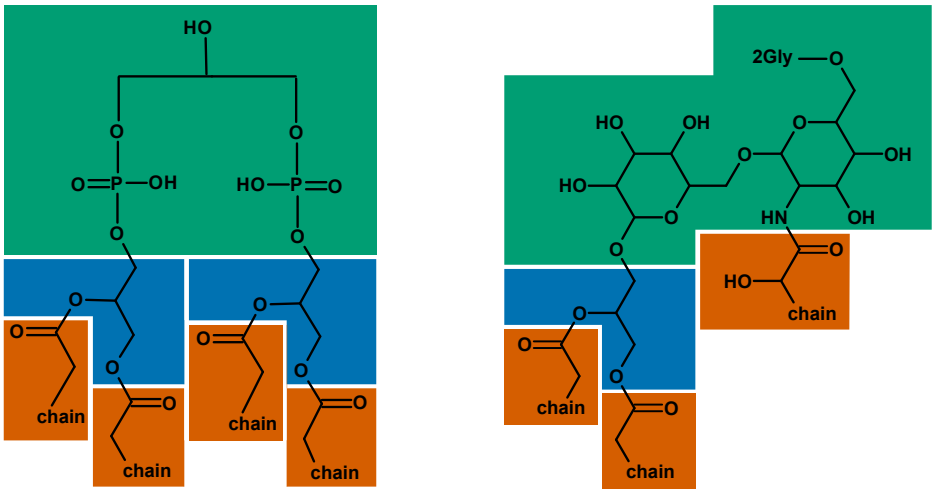
(g) GDGT



(h) 1,2-alkanediol



(i) NAcG-P-DAG



(j) DPG (k) 2GNAcG-G-DAG**

Figure 2.1: Structural designations used for IPL headgroups, backbones, and alkyl chains, for the sake of calculating abundance-weighted average properties and chemical formulas. Abbreviations are defined in the text. Green boxes contain structures associated with headgroups. For structures a through h, only the headgroup-backbone connector group explicitly shown, with the rest of the headgroup represented as 'head'. Putative full headgroup structures are shown for i through k. Blue boxes contain backbone structures. Orange boxes designate structures belonging to alkyl chains. Only the chemical structure of the first two carbons of each alkyl chain are shown; with 'chain' representing the rest. *In FAHFAm-OH, backbone-alkyl chain esterification may occur on either the 2' or 3' hydroxyl group (Diercks *et al.*, 2015). **Referred to as NAcG-DAG in Schubotz *et al.* (2013).

sugar headgroup in a glycolipid. Furthermore, the chemical formulae of all headgroups are calculated for their neutrally charged state, with the +1 charge imparted by quaternary ammonium functional of cationic lipids PC, TM-OL, TM-OL-OH, and TM-KL serving as the only exception because this charge is not the result of pH-dependent ionization.

Backbone structures are designated by blue boxes in Figure 2.1. In this study, IPL backbones were structurally designated according to three criteria chosen to promote consistency between observed IPL structures. First, the backbone must have a linear aliphatic chain of three carbons. Second, one carbon must be covalently bonded to an IPL headgroup. Third, the backbone must include two 'connector' functional groups that serve to anchor alkyl chains. Various backbone-alkyl chain linkage types are shown in Figure 2.2, with backbone connector groups shown proximal to the R_1 group representing the rest of the backbone; -CH_2- for a carbon-carbon (C-C) link (*e.g.*, Figure 2.2a), -O- for an ether link (*e.g.*, Figure 2.2b, c, d, and e), -NH- for an amide link (*e.g.*, Figure 2.2f), or -O- for an ester link (*e.g.*, Figure 2.2g, h, i, and j). The decision to include connector groups in the structure of the backbone rather than in the alkyl chain was made to maintain consistency when calculating nC in alkyl chains with C-C backbone-chain linkage relative to other chains. To demonstrate this point, consider the C-C linked alkyl chain and the ether-linked alkyl chain in Figure 2.2a and b. Both are saturated and have approximately the same physical length. To ensure that both chains have the same value for nC, the 'connector' groups proximal to R_1 must be categorized as part of the backbone structure.

In addition to these two connector groups, the three-carbon backbone may include modifications like hydroxylations or carbonyl groups. A single lipid may have more than one backbone, such as DPG or GDGT (Figure 2.1j and g).

Alkyl chains are aliphatic hydrocarbon chains linked to the IPL backbone or in

some cases, directly to the headgroup. Their structural designation is indicated by orange boxes in Figure 2.1. To calculate consistent nC across all observed variations in IPL structure, each chain begins at the carbon directly after the backbone 'connector' group and continues to the distal methyl group that caps the end of the chain. Alkyl chains of GDGTs provide the only exception to this, as they have two continuous monolayer-forming (membrane-spanning) isoprenoidal chains rather than two bilayer-forming chains common to non-GDGT IPLs. To calculate and compare chain properties across monolayer and bilayer-forming IPLs, each mole of GDGT was counted as having four moles of equal-length monolayer-forming chains (see Figure 2.1g), with each chain covalently bonded to one other chain by a -CH_2- group (*e.g.*, Figure 2.2d or e).

2.3.7 Calculation of average lipid properties and elemental composition

Abundance-weighted properties of IPL headgroups, backbones, and chains were calculated for a sample using the equation

$$\Xi = \frac{\sum_i \Xi_{ipl,i} \cdot x_i}{\sum_i n_{component,i} \cdot x_i}, \tag{2.2}$$

where Ξ indicates the average property of interest, $\Xi_{ipl,i}$ represents the property summed across all components in the i^{th} IPL with a mole fraction of x_i and $n_{component,i}$ number of the component of interest.

Abundance-weighted properties calculated in this way include nC, nUnsat, and the number of hydroxylations (nOH) per alkyl chain, the fraction of alkyl chains that form monolayers ($x_{monolayer}$), and the fraction backbone-alkyl chain linkage types with an ether (x_{ether}), ester (x_{ester}), amide (x_{amide}), or C-C (x_{C-C}) bond.

The abundance-weighted number of internal rings per GDGT (not per alkyl chain) in a sample was also calculated using Equation 2.2 by setting x_i to the mole fraction

of the i^{th} GDGT (rather than the i^{th} IPL), and setting $n_{component,i} \cdot x_i$ equal to 1, thereby producing a per-GDGT property and not a per-chain property.

Average chemical formulae of IPL components were determined from known charge and elemental abundance of carbon, hydrogen, nitrogen, oxygen, phosphorus, and sulfur atoms in the chemical structures as $\Xi_{ipl,i}$ in Equation 2.2 and then solving for the weighted average of each. The average chemical formulae of full IPL structures were calculated using the charge and elemental abundance summed across the i^{th} IPL while setting $n_{component,i} \cdot x_i$ equal to 1. Even when the structure of an IPL component was unclear (*e.g.*, the unidentified '223' headgroup observed in this study), its elemental composition was typically obtainable with accurate-mass spectrometry. This permitted the inclusion of ambiguous structures when calculating average elemental abundances.

2.3.8 Calculation of IPL Z_C

Z_C for IPLs and their component parts can be calculated using

$$Z_C = \frac{2o + 3n - 5p - 4s - h - Z}{c}, \tag{2.3}$$

where Z stands for the net charge and c, h, n, o, p, and s represent the number of atoms of carbon, hydrogen, nitrogen, oxygen, phosphorus, and sulfur in the chemical formula of interest. Hydrogen, oxygen, and nitrogen were assigned oxidation states of +1, -2, and -3. Sulfur within the sulfonic acid group of SQ-DAG was assigned an oxidation state of +4. Phosphorus was assigned an oxidation state of +5 to be consistent with that of phosphorus within the phosphate ion. Charge gained or lost by pH-dependent protonation or deprotonation, as is common in many lipid headgroups, does not require extra consideration, as this does not affect Z_C.

Equation 2.3 was used to determine the Z_C values of individual lipid structures,

such as those reported in Figure 2.2. It was also used to calculate abundance-weighted Z_C values in each sample; in this case, using the abundance-weighted charge and carbon, hydrogen, nitrogen, oxygen, phosphorus, and sulfur atoms in the average chemical formulae of IPLs and their components.

2.3.9 Simulation of analytical uncertainty

A Monte Carlo-style bootstrap sensitivity analysis was performed in which manually integrated IPL HPLC-MS peak areas were allowed to vary randomly by up to 30% of their original value over the course of 999 iterations. In addition, the analytical response factors applied to any headgroup-backbone combination listed in Table 2.1 were allowed to vary by up two orders of magnitude higher and lower. Abundance-weighted Z_C for lipids and their components were re-calculated from abundance-weighted average chemical formulae after each iteration.

2.4 Results

2.4.1 Water chemistry

Temperature, pH, and conductivity measurements corresponding to samples along the four hot spring outflow channels are shown in Table 2.2. As water flows from the source and cools, microbial communities in alkaline hot springs change spatially along the channel, with chemotrophs dominating the higher-temperature end and green-orange pigmented phototrophic cyanobacteria typically appearing at the lower-temperature end. All four outflow channels show downstream changes in water chemistry (indicated in Table 2.3), marked by an overall decrease in concentration of reduced inorganic dissolved species and an increase in the concentration of oxidized inorganic solutes.

Table 2.2: Selected geochemical and physical data from each sample site

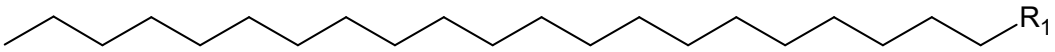
		12T UTM coordinates		Dist[a]	Zone[b]	Temperature	pH	Conductivity[c]
Site	Sample	easting	northing	(m)		(°C)		(μS/cm)
Bison Pool	BP1	510710	4935155	2.9	C	89.0	7.23	1550
	BP2	510715	4935156	8.2	C	80.9	7.34	1568
	BP3	510718	4935157	11.1	T	73.3	7.27	1540
	BP4	510719	4935159	13.4	P	63.1	8.09	-[d]
	BP5	510719	4935163	17.2	P	40.5	8.25	1508
	BP6	510724	4935165	22.6	P	29.0	9.01	1697
Mound Spring	MS1	511114	4934621	3.6	C	91.0	8.81	1612
	MS2	511108	4934624	12.7	C	77.3	8.65	1621
	MS3	511098	4934628	24.2	P	64.8	9.08	1617
	MS4	511083	4934621	38.7	P	53.0	9.22	1634
	MS5	511049	4934625	53	P	35.1	9.53	1660
Empress Pool	EP1	0521589	4948280	2.2	C	82.2	5.78	1824
	EP2	0521585	4948280	6.2	T	70.5	6.96	1832
	EP3	0521580	4948285	13.3	T	60.7	7.63	1840
	EP4	0521560	4948293	34.8	P	51.6	7.99	1860
	EP5	0521558	4948295	37.6	P	38.1	8.42	1664
Octopus Spring	OS1	0516054	4931217	7.0	C	85.4	7.29	1622
	OS2	0516016	4931212	38.3	P	59.8	8.27	1581

[a] Distance from hot spring source.

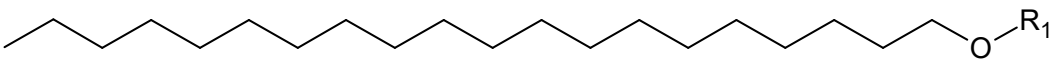
[b] Major metabolic regime representative of the microbial community at the sample site, interpreted visually in the field based on the presence or absence of photosynthetic pigments; C, strictly chemosynthetic; T, transition to phototrophy; P, photosynthetic.

[c] Conductivity was normalized to 25°C using the formula $\mathrm{Cond}_T/(1 + \alpha(T - 25))$, where Cond_T stands for the conductivity measured at the temperature of the sample site and α represents the temperature correction coefficient taken as 0.02 for freshwater.

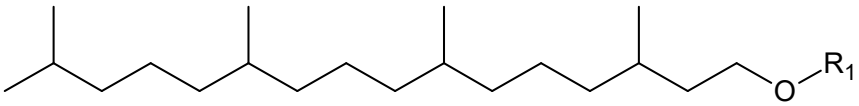
[d] No data.



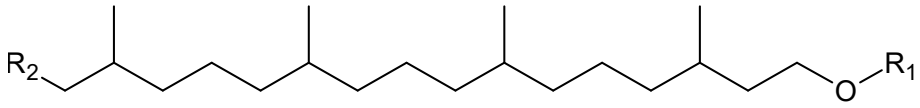
(a) C20:0 C-C-linked chain, $C_{21}H_{43}$, Z_C = -2.05



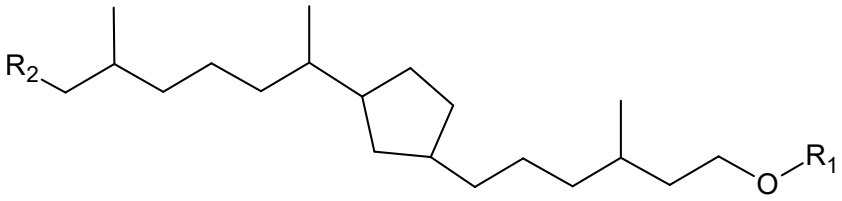
(b) C20:0 ether-linked non-isoprenoidal chain, $C_{20}H_{41}O$, Z_C = -1.95



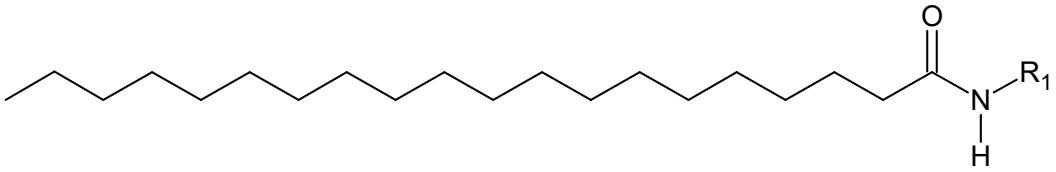
(c) C20:0 ether-linked isoprenoidal chain, $C_{20}H_{41}O$, Z_C = -1.95



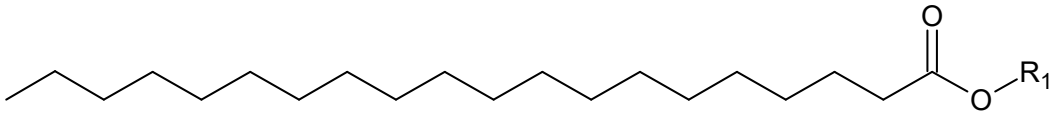
(d) C20:0 ether-linked isoprenoidal GDGT chain, $C_{20}H_{40}O$, Z_C = -1.90



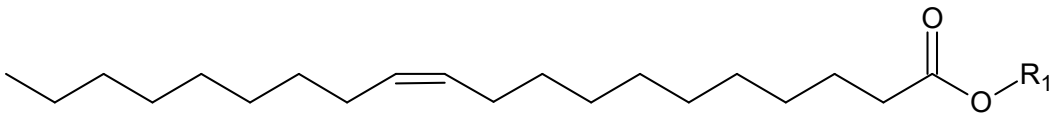
(e) C20:0 ether-linked isoprenoidal GDGT chain with internal ring, $C_{20}H_{38}O$, Z_C = -1.80



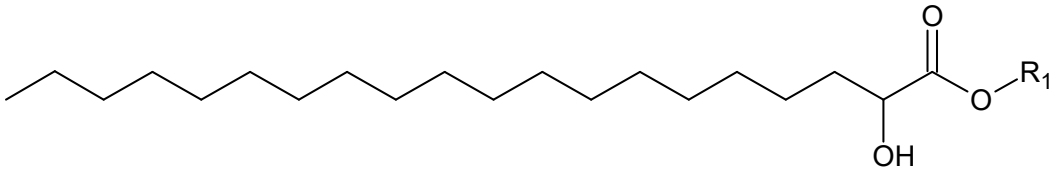
(f) C20:0 amide-linked chain, $C_{20}H_{40}NO$, Z_C = -1.75



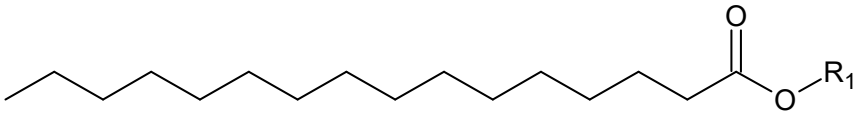
(g) C20:0 ester-linked fatty acid chain, $C_{20}H_{39}O_2$, Z_C = -1.75



(h) C20:1 ester-linked unsaturated fatty acid chain, $C_{20}H_{37}O_2$, Z_C = -1.65



(i) C20:0 ester-linked hydroxylated fatty acid chain, $C_{20}H_{39}O_3$, Z_C = -1.65



(j) C16:0 ester-linked fatty acid chain, $C_{16}H_{31}O_2$, Z_C = -1.35

Figure 2.2: Lipid alkyl chain modifications and backbone-chain linkage types organized by Z_C from reduced (top) to oxidized (bottom). Example structures were chosen to permit comparison of Z_C according to a various types of biochemical modifications: chain-backbone linkage type as C-C, ether, amide, or ester (a, b, f, g); non-branching and branching chains (b, c); bilayer- and monolayer-forming isoprenoidal chains (c, d); GDGT chains without and with an internal ring (d, e); saturated and unsaturated chains (g, h); non-hydroxylated and hydroxylated chains (g, i); and chains with a greater and lesser number of aliphatic carbons (g, j). R_1 represents the rest of the lipid backbone, and R_2 indicates the second half of a monolayer-forming GDGT chain and was considered a separate chain when calculating average chain properties.

Table 2.3: Concentrations of selected redox-sensitive dissolved chemical species[a]

		Oxidized				Reduced	
Site	Sample	O_2 (mg l^{-1})	NO_3^- (mg l^{-1})	NO_2^- (mg l^{-1})	$\sum SO_4^{2-}$ (mg l^{-1})	$\sum NH_4^+$ (mg l^{-1})	$\sum HS^-$ (μg l^{-1})
Bison Pool	BP1	0.2	0.01	0.02	13.11	0.07	230
	BP2	0.7	0.01	0.04	15.43	0.06	220
	BP3	1.1	0.02	0.01	16.81	0.04	bdl[b]
	BP4	2.3	0.03	0.02	16.50	0.02	6
	BP5	5.7	0.004	bdl	17.18	0.01	15
	BP6	3.3	0.07	bdl	18.32	0.02	10
Mound Spring	MS1	0.4	0.01	bdl	14.33	0.07	716
	MS2	2.2	0.01	bdl	15.03	0.01	758
	MS3	1.4	0.04	0.02	16.99	0.03	236
	MS4	3.6	0.02	0.01	17.56	0.02	70
	MS5	6.9	0.06	bdl	20.11	bdl	bdl
Empress Pool	EP1	0.4	0.01	0.08	106.87	0.42	260
	EP2	0.7	-[c]	-	-	-	97
	EP3	1.2	0.01	bdl	106.70	0.31	37
	EP4	1.3	0.03	0.03	111.70	0.39	31
	EP5	3.4	0.08	0.01	111.24	0.14	18
Octopus Spring	OS1	0.5	0.03	0.03	17.82	0.06	13
	OS2	3.3	0.03	0.02	18.76	0.02	12

[a] Note: Sulfide (HS^-), ammonium (NH_4^+), and sulfate (SO_4^{2-}) concentrations are summed for their respective pH-dependent protonated states.
[b] bdl: below detection limit.
[c] No data.

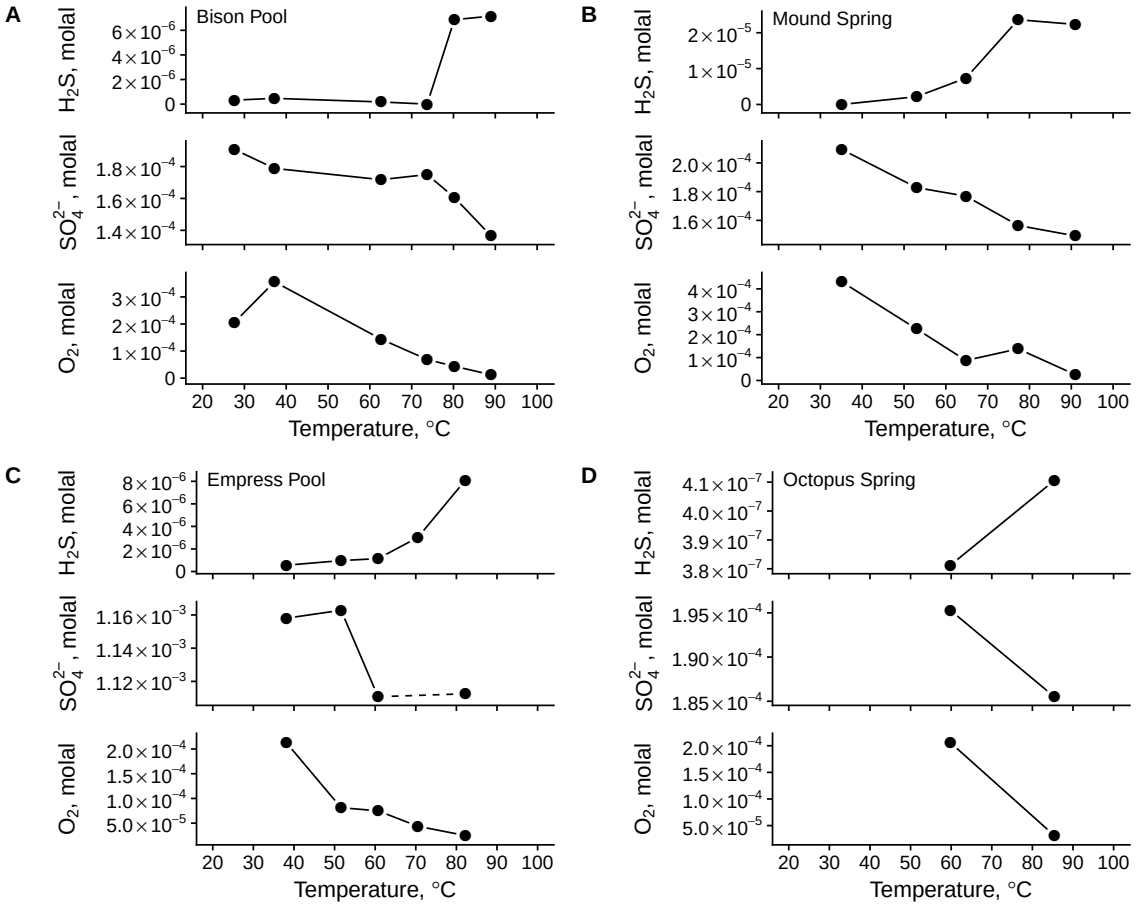


Figure 2.3: Total concentrations of redox-sensitive aqueous chemical species in samples from Bison Pool (A), Mound Spring (B), Empress Pool (C) and Octopus Spring (D). Lines between points are meant to guide the eye between measurements only. A water sample was not collected for sulfate at Empress Pool site EP2 during the 2012 field season, indicated here by a dashed line between sulfate measurements for sites EP1 and EP3.

It can be seen from Figure 2.3 that concentrations of relatively reduced sulfide are higher at the hot spring source and gradually give way to sulfate, likely due to biological oxidation (Cox *et al.*, 2011). Previous research at Bison Pool and Mound Spring by Loiacono *et al.* (2012) demonstrated active expression of nitrogen-metabolizing genes by nitrifying microorganisms, possibly explaining the downstream decrease of ammonia concentrations and a concurrent increase in nitrite and nitrate concentrations reported in Table 2.3. The dissolved oxygen also exhibits an overall increase in concentration downstream that is likely attributable to mixing with atmospheric O_2 and input from oxygenic photosynthesis from cyanobacteria. Collectively, these geochemical data suggest the redox potential of the water was generally more reduced at the hot spring source and gradually becomes more oxidized downstream.

2.4.2 Z_C of IPLs and their components

The abundance-weighted chemical formulae and Z_C calculated for thermophilic microbial IPLs and their headgroups, backbones, and alkyl chains are reported in Table 2.4 for samples collected from Bison Pool, Mound Spring, Empress Pool, and Octopus Spring. As shown in Figure 2.4, decreasing temperature and increasingly oxidized conditions coincide with near-linear changes in the weighted Z_C of IPLs and their components. Carbon was most reduced in IPLs sampled closest to the spring sources (Z_C between -1.68 and -1.56) where temperatures were highest (82.2 to 91.0°C) and concentrations of oxidized inorganic species (nitrate, sulfate, and oxygen) were lowest. In progressively downstream samples, carbon in IPLs became more oxidized, with weighted Z_C between -1.36 and -1.33 for samples in the 29.0 to 38.1°C temperature range where measured concentrations of sulfate, nitrate, and oxygen tended to be greatest. This trend persists even after simulating potential sources of analytical error, as shown in Figure 2.4 by the black bars above and below the 'canon'

Table 2.4: Weighted Z_C and chemical formula of IPLs and their component parts

		Weighted Z_C				Weighted chemical formula			
Site	Sample	Full IPLs	Headgroups	Backbones	Alkyl chains	Full IPLs	Headgroups	Backbones	Alkyl chains
Bison	BP1	-1.56	0.15	-0.34	-1.92	$C_{64.1}H_{123.}N_{2.15e\text{-}1}O_{12.3}P_{5.01e\text{-}1}S_{1.27e\text{-}3}{}^{+8.04e\text{-}3}$	$C_{5.61}H_{11.0}N_{1.46e\text{-}1}O_{6.70}P_{3.95e\text{-}1}S_{9.55e\text{-}4}{}^{+6.07e\text{-}3}$	$C_{3.02}H_{5.05}N_{1.71e\text{-}2}O_{1.99}$	$C_{19.8}H_{38.5}O_{2.85e\text{-}1}$
Pool	BP2	-1.53	0.15	-0.34	-1.89	$C_{56.9}H_{109.}N_{2.10e\text{-}1}O_{12.3}P_{6.07e\text{-}1}S_{2.42e\text{-}3}{}^{+1.42e\text{-}2}$	$C_{6.25}H_{12.1}N_{1.58e\text{-}1}O_{7.62}P_{5.34e\text{-}1}S_{2.05e\text{-}3}{}^{+1.21e\text{-}2}$	$C_{3.04}H_{5.07}N_{2.25e\text{-}2}O_{1.99}$	$C_{19.4}H_{37.6}O_{4.36e\text{-}1}$
	BP3	-1.45	0.11	-0.35	-1.84	$C_{48.0}H_{91.5}N_{1.48e\text{-}1}O_{12.1}P_{5.31e\text{-}1}S_{8.32e\text{-}2}{}^{+4.61e\text{-}3}$	$C_{6.71}H_{12.9}N_{1.03e\text{-}1}O_{8.10}P_{5.12e\text{-}1}S_{7.95e\text{-}2}{}^{+4.40e\text{-}3}$	$C_{3.06}H_{5.12}N_{3.89e\text{-}2}O_{1.97}$	$C_{17.9}H_{34.5}O_{7.54e\text{-}1}$
	BP4	-1.46	0.13	-0.35	-1.88	$C_{45.1}H_{86.8}N_{8.39e\text{-}2}O_{12.1}P_{6.25e\text{-}1}S_{5.38e\text{-}2}{}^{+8.07e\text{-}3}$	$C_{7.22}H_{13.7}N_{5.44e\text{-}2}O_{8.90}P_{6.23e\text{-}1}S_{5.36e\text{-}2}{}^{+8.04e\text{-}3}$	$C_{3.04}H_{5.09}N_{2.91e\text{-}2}O_{1.97}$	$C_{17.3}H_{33.7}O_{5.67e\text{-}1}$
	BP5	-1.38	0.04	-0.32	-1.80	$C_{44.6}H_{83.7}N_{3.32e\text{-}1}O_{11.4}P_{2.17e\text{-}1}S_{1.21e\text{-}1}{}^{+1.24e\text{-}1}$	$C_{7.82}H_{14.6}N_{3.26e\text{-}1}O_{7.69}P_{2.17e\text{-}1}S_{1.21e\text{-}1}{}^{+1.24e\text{-}1}$	$C_{3.11}H_{5.02}N_{5.65e\text{-}3}O_{2.01}$	$C_{16.8}H_{32.0}O_{8.49e\text{-}1}$
	BP6	-1.36	0.02	-0.33	-1.73	$C_{43.5}H_{80.4}N_{2.70e\text{-}1}O_{11.6}P_{4.64e\text{-}1}S_{1.18e\text{-}1}{}^{+8.39e\text{-}2}$	$C_{6.67}H_{13.1}N_{2.63e\text{-}1}O_{7.57}P_{4.63e\text{-}1}S_{1.18e\text{-}1}{}^{+8.37e\text{-}2}$	$C_{3.04}H_{5.02}N_{6.73e\text{-}3}O_{2.00}$	$C_{16.7}H_{30.8}O_{9.75e\text{-}1}$
Mound	MS1	-1.68	0.32	-0.33	-1.94	$C_{92.2}H_{177.}N_{4.49e\text{-}4}O_{11.2}P_{2.43e\text{-}4}{}^{+7.34e\text{-}6}$	$C_{3.14}H_{6.23}N_{1.29e\text{-}4}O_{3.61}P_{1.22e\text{-}4}{}^{+3.67e\text{-}6}$	$C_{3.00}H_{5.00}N_{9.56e\text{-}5}O_{2.00}$	$C_{20.0}H_{38.8}O_{3.52e\text{-}4}$
Spring	MS2	-1.61	0.18	-0.38	-1.93	$C_{68.7}H_{133.}N_{2.22e\text{-}1}O_{12.0}P_{3.77e\text{-}1}S_{2.10e\text{-}2}{}^{+4.40e\text{-}3}$	$C_{5.12}H_{9.87}N_{7.51e\text{-}2}O_{5.98}P_{2.65e\text{-}1}S_{1.46e\text{-}2}{}^{+3.06e\text{-}3}$	$C_{3.09}H_{5.24}N_{8.02e\text{-}2}O_{1.92}$	$C_{19.6}H_{38.2}O_{2.12e\text{-}1}$
	MS3	-1.52	0.14	-0.35	-1.92	$C_{49.9}H_{96.5}N_{4.98e\text{-}2}O_{11.9}P_{6.47e\text{-}1}S_{2.03e\text{-}2}{}^{+1.27e\text{-}3}$	$C_{6.57}H_{12.6}N_{1.72e\text{-}2}O_{8.26}P_{5.94e\text{-}1}S_{1.86e\text{-}2}{}^{+1.17e\text{-}3}$	$C_{3.03}H_{5.09}N_{2.84e\text{-}2}O_{1.97}$	$C_{18.0}H_{35.3}O_{3.30e\text{-}1}$
	MS4	-1.43	0.06	-0.34	-1.82	$C_{47.2}H_{89.6}N_{2.07e\text{-}1}O_{12.1}P_{4.93e\text{-}1}S_{5.34e\text{-}2}{}^{+5.00e\text{-}2}$	$C_{7.20}H_{13.8}N_{1.91e\text{-}1}O_{8.12}P_{4.78e\text{-}1}S_{5.18e\text{-}2}{}^{+4.85e\text{-}2}$	$C_{3.03}H_{5.03}N_{9.59e\text{-}3}O_{1.99}$	$C_{17.4}H_{33.3}O_{7.89e\text{-}1}$
	MS5	-1.33	-0.07	-0.34	-1.71	$C_{46.6}H_{85.5}N_{2.60e\text{-}1}O_{12.5}P_{3.65e\text{-}1}S_{1.25e\text{-}1}{}^{+1.69e\text{-}1}$	$C_{8.29}H_{15.9}N_{2.51e\text{-}1}O_{8.36}P_{3.62e\text{-}1}S_{1.24e\text{-}1}{}^{+1.68e\text{-}1}$	$C_{3.04}H_{5.02}N_{7.05e\text{-}3}O_{1.99}$	$C_{16.8}H_{30.7}O_{9.51e\text{-}1}$
Empress	EP1	-1.60	0.22	-0.33	-1.96	$C_{85.0}H_{164.}N_{1.17e\text{-}1}O_{13.8}P_{4.38e\text{-}2}{}^{+2.05e\text{-}4}$	$C_{5.66}H_{10.4}N_{6.51e\text{-}2}O_{5.79}P_{2.97e\text{-}2}{}^{+1.16e\text{-}4}$	$C_{3.00}H_{5.01}N_{2.23e\text{-}3}O_{2.00}$	$C_{19.8}H_{38.8}O_{1.61e\text{-}2}$
Pool	EP2	-1.54	0.15	-0.36	-1.92	$C_{62.0}H_{119.}N_{1.87e\text{-}1}O_{12.5}P_{4.18e\text{-}1}S_{1.16e\text{-}2}{}^{+8.26e\text{-}3}$	$C_{6.22}H_{11.8}N_{8.79e\text{-}2}O_{7.06}P_{3.28e\text{-}1}S_{8.81e\text{-}3}{}^{+6.30e\text{-}3}$	$C_{3.07}H_{5.17}N_{5.52e\text{-}2}O_{1.95}$	$C_{18.9}H_{36.8}O_{2.61e\text{-}1}$
	EP3	-1.54	0.15	-0.34	-1.90	$C_{67.9}H_{130.}N_{1.06e\text{-}1}O_{13.4}P_{3.20e\text{-}1}S_{3.36e\text{-}2}{}^{+1.15e\text{-}2}$	$C_{6.11}H_{11.5}N_{5.37e\text{-}2}O_{6.74}P_{2.32e\text{-}1}S_{2.34e\text{-}2}{}^{+7.98e\text{-}3}$	$C_{3.03}H_{5.06}N_{2.05e\text{-}2}O_{1.98}$	$C_{19.1}H_{36.9}O_{2.95e\text{-}1}$
	EP4	-1.44	0.08	-0.33	-1.83	$C_{51.6}H_{97.2}N_{2.52e\text{-}1}O_{11.7}P_{1.99e\text{-}1}S_{1.13e\text{-}1}{}^{+8.90e\text{-}2}$	$C_{6.88}H_{12.9}N_{2.11e\text{-}1}O_{7.00}P_{1.78e\text{-}1}S_{9.96e\text{-}2}{}^{+7.87e\text{-}2}$	$C_{3.07}H_{5.04}N_{1.18e\text{-}2}O_{2.00}$	$C_{17.8}H_{33.9}O_{6.78e\text{-}1}$
	EP5	-1.37	-0.01	-0.31	-1.77	$C_{48.1}H_{89.0}N_{3.69e\text{-}1}O_{11.6}P_{8.99e\text{-}2}S_{1.79e\text{-}1}{}^{+1.57e\text{-}1}$	$C_{7.78}H_{14.6}N_{3.44e\text{-}1}O_{7.23}P_{8.53e\text{-}2}S_{1.68e\text{-}1}{}^{+1.47e\text{-}1}$	$C_{3.11}H_{5.00}N_{1.48e\text{-}3}O_{2.01}$	$C_{17.1}H_{31.9}O_{8.12e\text{-}1}$
Octopus	OS1	-1.56	0.13	-0.34	-1.95	$C_{64.4}H_{125.}N_{9.12e\text{-}2}O_{13.7}P_{5.91e\text{-}1}S_{2.83e\text{-}3}{}^{+1.45e\text{-}3}$	$C_{6.97}H_{13.3}N_{8.14e\text{-}2}O_{8.32}P_{5.36e\text{-}1}S_{2.23e\text{-}3}{}^{+1.15e\text{-}3}$	$C_{3.01}H_{5.02}N_{5.92e\text{-}3}O_{1.99}$	$C_{20.4}H_{40.3}O_{2.63e\text{-}1}$
Spring	OS2	-1.48	0.12	-0.34	-1.89	$C_{45.2}H_{87.3}N_{1.01e\text{-}1}O_{11.9}P_{6.79e\text{-}1}S_{4.19e\text{-}2}{}^{+1.93e\text{-}2}$	$C_{6.91}H_{13.3}N_{8.39e\text{-}2}O_{8.72}P_{6.78e\text{-}1}S_{4.18e\text{-}2}{}^{+1.93e\text{-}2}$	$C_{3.03}H_{5.05}N_{1.69e\text{-}2}O_{1.99}$	$C_{17.5}H_{34.2}O_{5.73e\text{-}1}$

Note: fractional charge is represented by a + followed by its fraction.

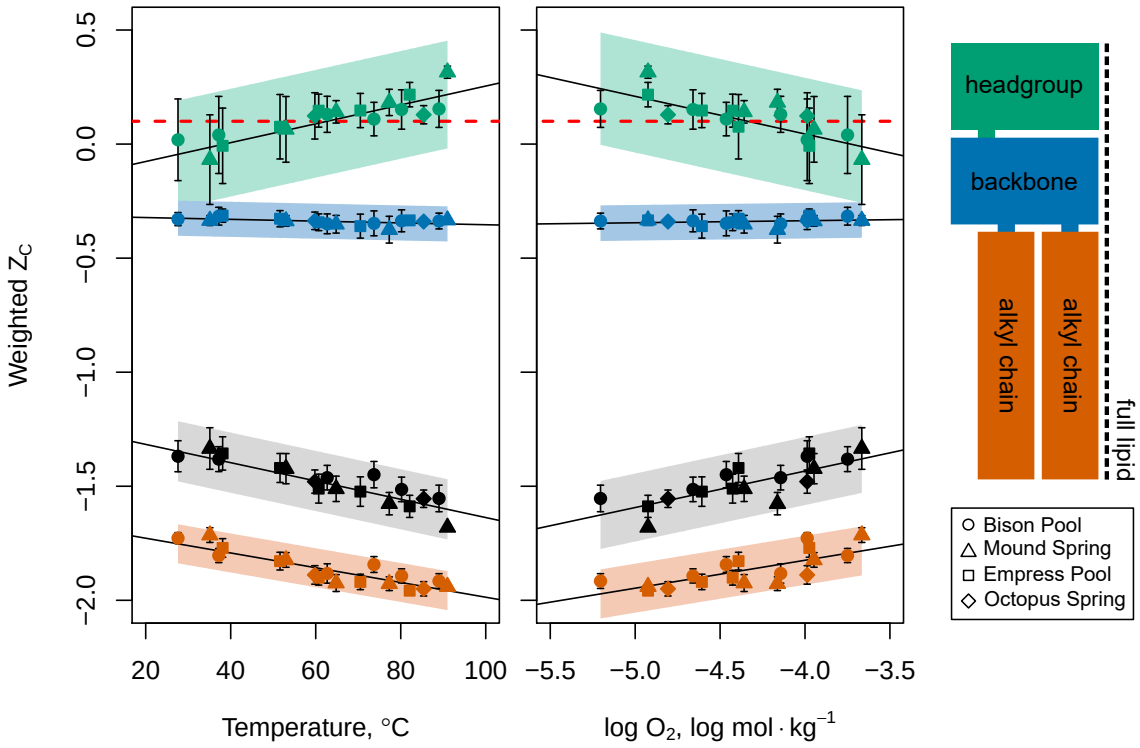


Figure 2.4: Abundance-weighted Z_C of headgroups (green), backbones (blue), alkyl chains (orange) and full (black) thermophile IPLs sampled along the outflow channels of four Yellowstone hot springs with respect to temperature (left) and log molality of dissolved O_2 (right). Symbols designate the observed values of weighted Z_C of extracted lipids and their components. Bars around the points show the standard deviation of 999 weighted Z_C resulting from the bootstrap sensitivity analysis. Linear regressions of these bootstrap values are indicated by best-fit lines. Shaded areas represent the predicted 95% intervals for the Z_C values produced by sensitivity analysis. The dashed red line at a weighted Z_C of 0.1 indicates one possible value encompassed by most of the standard deviation bars produced by the sensitivity analysis for IPL headgroups.

observations of abundance-weighted Z_C. These bars indicate the standard deviation of weighted Z_C resulting from 999 iterations of the bootstrap sensitivity analysis. The standard deviations of IPL weighted Z_C show little overlap between upstream and downstream samples after random variation of up to 30% for integrated HPLC-MS peak areas and up to two orders of magnitude for response factors.

Alkyl chains tend to have weighted Z_C approximately 0.3-0.4 more negative than those of the full structure regardless of the sample. Compared to headgroups and backbones, the weighted Z_C of alkyl chains were closest to that of the full structure, implicating this component as the primary contributor to full IPL Z_C and the observed trends with temperature and redox state. This agrees with the observation that most carbon in IPLs belongs to the alkyl chains. Since alkyl chains have a high H/C ratio to facilitate their hydrophobicity, the weighted Z_C of full IPLs and their chains are significantly more negative than those of the backbones and headgroups with a higher ratio of electronegative atoms (*e.g.*, oxygen, nitrogen) to carbon. As shown in Figure 2.4, trends in alkyl chain Z_C appear to be resistant to simulated sources of analytical uncertainty than the full structure based on sample standard deviations and the narrow 95% prediction interval generated by the bootstrap sensitivity analysis. The width of the prediction interval for future bootstrap results does not overlap between the highest temperature and lowest temperature, which I interpret as further evidence that the downstream increase in alkyl chain Z_C is not an artifact of the analytical method used to quantify lipid abundance.

The weighted Z_C of IPL backbones did not change significantly with temperature or redox state, as most observed backbones were composed of glycerol regardless of the sample location. According to the IPL component division scheme used in this study, an IPL glycerol backbone has a chemical formula of $C_3H_5O_2$, corresponding to a Z_C of $-0.\overline{3}$. This was close to the abundance-weighted value in every sample, with

little variation.

Carbon in headgroups was significantly more oxidized than in any other component, with weighted Z_C ranging from -0.01 to 0.32 across samples. Glycolipid headgroups 1G ($Z_C = 0.1\overline{6}$) and 2G ($Z_C = 0.08\overline{3}$) were among the most abundant headgroups observed in all four springs regardless of temperature or redox state, though they were most abundant, along with the glycolipid SQ ($Z_C = 0.1\overline{6}$) in samples where photosynthetic microorganisms are visually apparent. This agrees with observations that 1G, 2G, and SQ glycolipids are abundant in cyanobacterial (Wada and Murata, 2009) and algal (Guschina and Harwood, 2006) photosynthetic membranes.

Phosphate-bearing phospholipids and glycophospholipids, especially lipids with PI headgroups, were more abundant downstream than upstream. However, the incorporation of phosphate into the glycolipid headgroups had little to no effect on abundance-weighted headgroup Z_C, as there is no difference in Z_C between analogous nonphosphorylated and phosphorylated headgroups (*e.g.*, between 1G-P and 1G, or between 2G-P and 2G, *etc.*). The greatest abundance of lipids with an APT phospholipid headgroup (Z_C = -0.2) was observed in 'pink streamer' thermophile communities sampled from the upstream chemosynthetic zones of Bison Pool and Octopus Spring; this result agrees with previous reports that this lipid is common in streamers dominated by *Aquificales* bacteria (Sturt *et al.*, 2004; Schubotz *et al.*, 2013). However, after applying response factors, the abundance APT was low relative to other headgroups, even in streamer samples.

Intriguingly, weighted Z_C of IPL headgroups became progressively negative with decreasing temperature or increasingly oxidized conditions, the reverse of trends observed alkyl chains and full IPLs. One reason for this stems from the way abundance-weighted Z_C was calculated for GDGT headgroups. Because of the monolayer-forming properties of GDGTs, one mole of GDGT will have two moles of headgroup. To il-

lustrate, consider that 1G-GDGT has two headgroups; the first is a glycosyl moiety linked to glycerol on one side of the GDGT, and the second is a hydroxyl (-OH) group terminating the glycerol on the other side of the GDGT. Summing the elemental abundances of 1G ($C_6H_{11}O_6$) and -OH results in a total of $C_6H_{12}O_7$ for both headgroups of 1G-GDGT, or $C_3H_6O_{3.5}$ per headgroup. Accounting for this extra -OH means that carbon in the headgroups of 1G-GDGT ($Z_C = 0.\overline{3}$) is considerably more oxidized than carbon in 1G headgroups of non-GDGTs ($Z_C = 0.1\overline{6}$). While 1G served as an example here, the extra oxygen and hydrogen atom must be accounted for in *all* GDGTs, resulting in more oxidized headgroup carbon than one might expect based on headgroup identity in high-temperature GDGT-dominated samples. Another factor contributing to the observed trends in the weighted Z_C of IPL headgroups was a greater abundance of multi-methylated headgroups downstream, such as PC ($Z_C = -1.4$), TM-OL ($Z_C = -0.875$), TM-lysine ($Z_C = -1.00$). Each methylation increases the H/C ratio of a headgroup substantially, resulting in more reduced carbon.

Regardless, trends in the weighted Z_C of IPL headgroups appear to be substantially influenced by the sources of analytical uncertainty simulated by the bootstrap sensitivity analysis, especially for the low-temperature downstream samples. As indicated by the horizontal, dashed red line in Figure 2.4, weighted Z_C of headgroups in most samples show overlap in the vertical bars representing the simulated standard deviation at a value of about 0.1. Furthermore, the 95% prediction interval for future bootstrap calculations was widest for headgroups, with significant overlap between prediction range at the samples furthest upstream and downstream. This calls into question the significance of the apparent headgroup Z_C correlation with temperature and redox state, as these trends may be artifacts of the analytical method used to quantify lipids.

Table 2.5: Summary of abundance-weighted alkyl chain properties

Site	Sample	nC	nUnsat	nOH	Mole fraction alkyl chain or linkage type					Rings per GDGT
					$x_{monolayer}$	x_{ether}	x_{ester}	x_{amide}	x_{C-C}	
Bison Pool	BP1	19.80	3.12×10^{-1}	2.10×10^{-3}	4.90×10^{-1}	7.03×10^{-1}	2.75×10^{-1}	1.11×10^{-2}	1.09×10^{-2}	1.8
	BP2	19.39	3.62×10^{-1}	5.27×10^{-3}	3.00×10^{-1}	5.46×10^{-1}	4.18×10^{-1}	1.64×10^{-2}	1.92×10^{-2}	2.2
	BP3	17.94	3.34×10^{-1}	5.91×10^{-3}	8.81×10^{-2}	2.17×10^{-1}	7.32×10^{-1}	2.37×10^{-2}	2.79×10^{-2}	2.8
	BP4	17.28	3.46×10^{-1}	4.70×10^{-3}	7.16×10^{-3}	4.14×10^{-1}	5.52×10^{-1}	1.51×10^{-2}	1.88×10^{-2}	3.1
	BP5	16.82	4.91×10^{-1}	2.93×10^{-4}	7.20×10^{-4}	9.80×10^{-2}	8.46×10^{-1}	2.83×10^{-3}	5.33×10^{-2}	3.2
	BP6	16.68	8.10×10^{-1}	3.03×10^{-3}	2.94×10^{-3}	3.28×10^{-3}	9.72×10^{-1}	3.36×10^{-3}	2.15×10^{-2}	3.3
Mound Spring	MS1	20.00	2.72×10^{-4}	-	9.99×10^{-1}	< 1.00	3.04×10^{-4}	4.78×10^{-5}	4.96×10^{-5}	2.5
	MS2	19.60	1.94×10^{-1}	7.79×10^{-4}	6.07×10^{-1}	7.39×10^{-1}	1.71×10^{-1}	4.74×10^{-2}	4.33×10^{-2}	2.0
	MS3	18.01	3.53×10^{-1}	7.06×10^{-4}	1.64×10^{-1}	6.54×10^{-1}	3.15×10^{-1}	1.55×10^{-2}	1.50×10^{-2}	2.8
	MS4	17.38	3.88×10^{-1}	1.48×10^{-4}	5.85×10^{-2}	1.94×10^{-1}	7.84×10^{-1}	5.10×10^{-3}	1.69×10^{-2}	2.9
	MS5	16.78	9.65×10^{-1}	3.39×10^{-5}	1.09×10^{-2}	2.68×10^{-2}	9.47×10^{-1}	3.52×10^{-3}	2.23×10^{-2}	3.5
Empress Pool	EP1	19.83	1.51×10^{-2}	2.35×10^{-4}	8.63×10^{-1}	9.83×10^{-1}	1.47×10^{-2}	1.11×10^{-3}	1.13×10^{-3}	2.1
	EP2	18.91	2.24×10^{-1}	1.34×10^{-2}	4.71×10^{-1}	7.06×10^{-1}	2.30×10^{-1}	3.15×10^{-2}	3.25×10^{-2}	2.5
	EP3	19.06	1.37×10^{-1}	6.14×10^{-3}	6.06×10^{-1}	6.88×10^{-1}	2.85×10^{-1}	1.07×10^{-2}	1.65×10^{-2}	2.5
	EP4	17.78	3.97×10^{-1}	3.88×10^{-3}	2.31×10^{-1}	2.84×10^{-1}	6.72×10^{-1}	5.91×10^{-3}	3.75×10^{-2}	2.5
	EP5	17.08	6.87×10^{-1}	4.47×10^{-4}	1.26×10^{-1}	1.32×10^{-1}	8.11×10^{-1}	7.42×10^{-4}	5.62×10^{-2}	2.3
Octopus Spring	OS1	20.42	1.86×10^{-1}	2.50×10^{-4}	4.21×10^{-1}	7.38×10^{-1}	2.55×10^{-1}	3.24×10^{-3}	3.28×10^{-3}	1.1
	OS2	17.52	3.17×10^{-1}	5.41×10^{-3}	7.60×10^{-3}	4.10×10^{-1}	5.64×10^{-1}	8.71×10^{-3}	1.73×10^{-2}	2.3

2.4.3 Abundance-weighted alkyl chain properties

The overall increase in the oxidation state of full IPLs downstream along the studied outflow channels was primarily caused by a shift in abundance-weighted alkyl chain properties. These properties are reported in Table 2.5. Alkyl chain modifications with the most influence on weighted Z_C were nC, nUnsat, the number of internal pentacyclic rings per GDGT, and backbone-alkyl chain linkage chemistry. Other alkyl chain modifications, such as nOH or monolayer-forming characteristics, were observed at low abundance or did not substantially change lipid chemical formulae, and as a result, did not notably impact weighted Z_C.

The transition from hot, reduced conditions upstream to cool, oxidized conditions downstream was characterized by a gradual switch from ether- to ester-dominated IPL backbone-chain linkage, as shown in Figure 2.5. Upstream samples were rich in GDGT and DEG lipids containing exclusively ether-linked alkyl chains. Ether-linked alkyl chains are thought to be resistant to hydrolysis at high-temperature (Daniel and Cowan, 2000). Further from the hot spring source, within the chemosynthetic zone, abundances of AEG lipids became increasingly more prominent, each bearing one ether- and one ester-linked chain; although the abundance of these 'hybrid'-linked lipids was not observed to surpass GDGT, DEG, or DAG in any sample. The transition to downstream photosynthetic communities was associated with a sharp increase in exclusively ester-linked DAG lipids. Together, these changes in chain-backbone linkage chemistry correspond to a downstream increase in the Z_C of microbial IPLs stemming from the difference in ester and ether chemical formulae. An ester bond has one more oxygen and two fewer hydrogen atoms than an ether bond (compare structures in Figure 2.2b and g), resulting in more oxidized carbon in an ester bond relative to an ether bond. Amide and C-C linkage types were present

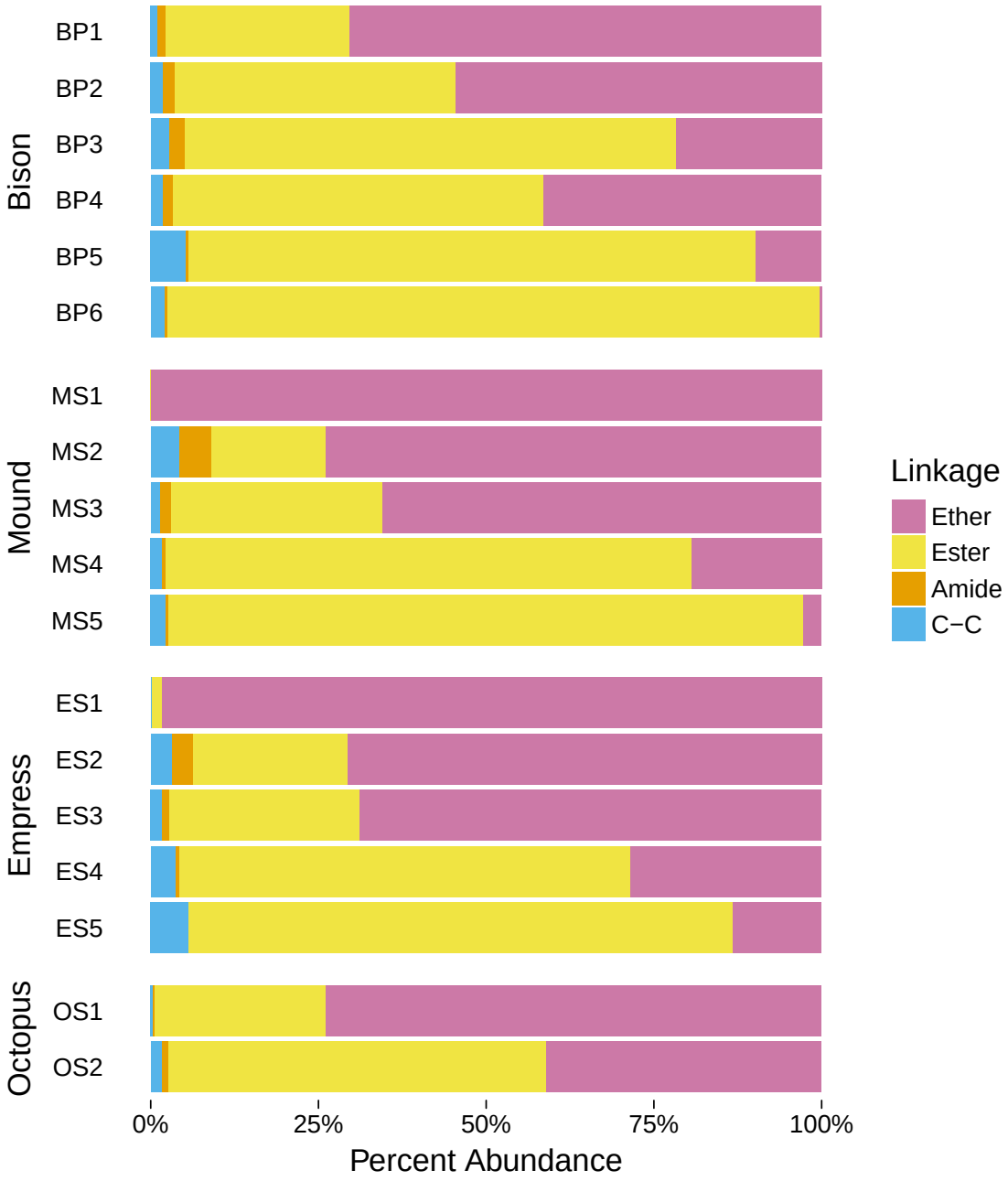


Figure 2.5: Relative abundance of backbone-alkyl chain linkage types in the four hot springs. Examples of C-C, ether, amide, and ester-linked alkyl chains are portrayed in Figure 2.2a, b, f, and g, respectively.

at relatively low abundance (typically less than 5%).

The weighted nC of alkyl chains decreased downstream in all four hot springs, as shown in Figure 2.6. This trend agrees with a plethora of studies demonstrating the capacity of microorganisms to adapt their membrane fluidity and permeability in response to temperature by adjusting the strength of hydrophobic interaction within the nonpolar portion of their membranes (see review by van Meer *et al.*, 2008). In the highest-temperature samples, weighted nC was approximately 20 since these samples were rich in GDGTs; each has 80 aliphatic carbons distributed across all alkyl chains. According to the structural division scheme used in this study, one mole of GDGT corresponds to four moles of equal-length alkyl chains, resulting in an nC of 20 per chain. Weighted nC was observed to exceed 20 in sample OS1 due to an abundance of long-chain archaeol lipids, with 50 or 55 aliphatic carbons split between two alkyl chains for an average of 25 and 27.5 carbons per chain.

Alkyl chains can be modified with unsaturations, or double bonds, in both the *cis* and *trans* configurations. A *cis*-unsaturation is thought to increase membrane fluidity by introducing a 'kink' in an alkyl chain that decreases chain packing and disrupts neighboring lipids in the membrane. A *trans*-unsaturation has been shown to increase growth temperature (Kiran *et al.*, 2005) and resistance to solvents and desiccation (Halverson and Firestone, 2000). An alkyl chain containing an unsaturation (Figure 2.2h) is more oxidized than a saturated alkyl chain (Figure 2.2g) because it has two fewer hydrogen atoms. The weighted number of unsaturations per alkyl chain increases downstream in all four outflow channels (Figure 2.7). This trend was most prominent at Mound Spring and Empress Pool, where the average degree of unsaturation was close to zero for samples nearer to the source and increases gradually to nearly one unsaturation per alkyl chain in samples furthest downstream.

The incorporation of one or more rings into the alkyl chains of GDGTs has been

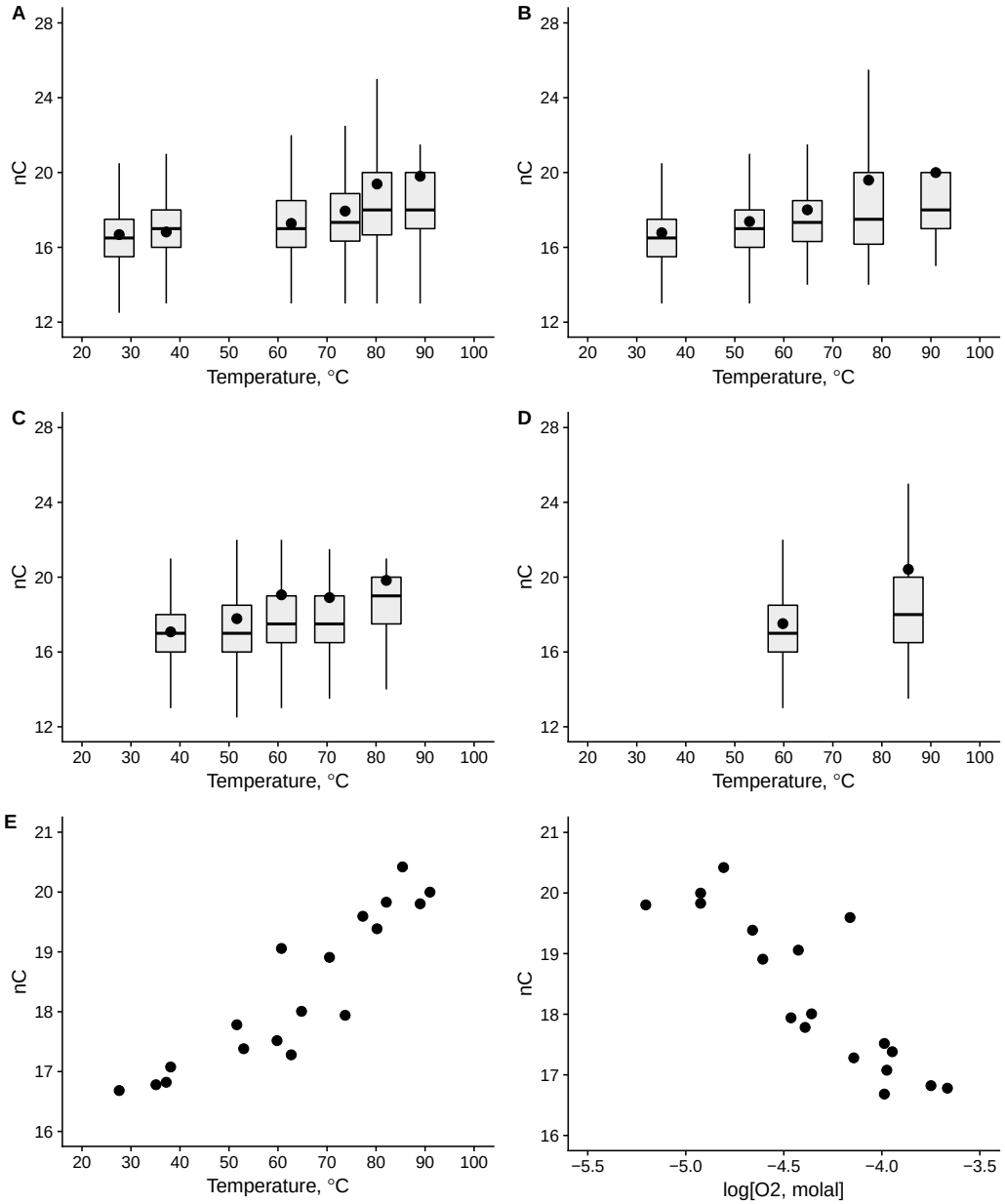


Figure 2.6: Number of aliphatic carbons (nC) per IPL alkyl chain as a function of temperature. Points are the average weighted value, rectangles show the interquartile range (IQR) of observed values, with 25% of observations lying above and below the black middle line, representing the median value, with whiskers encompassing observations up to 1.5 times the IQR beyond this range. Values that fall outside the range of the whiskers are not indicated. The abundance-weighted nC is shown for all samples in panel E as a function of temperature (left) and dissolved oxygen concentration (right).

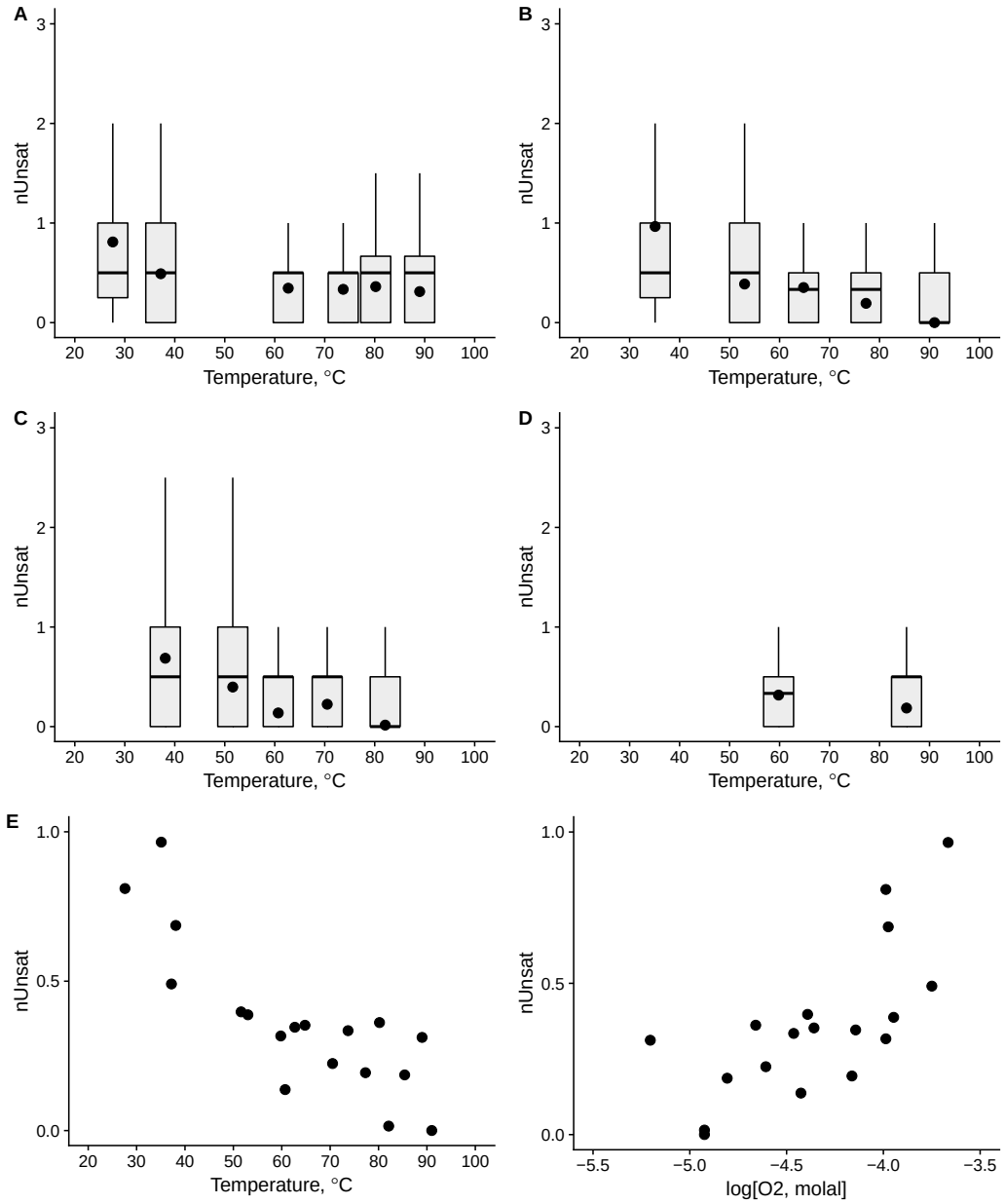


Figure 2.7: Number of unsaturations (nUnsat) per IPL alkyl chain. Refer to the caption of Figure 2.6 for a description of the symbols and statistics used in panels A-D. The abundance-weighted number of unsaturations per alkyl chain is shown for all samples in panel E as a function of temperature (left) and dissolved oxygen concentration (right).

proposed to enhance lipid packing while increasing fluidity in archaeal membranes (Sollich *et al.*, 2017). With regards to the influence of GDGT rings on weighted Z_C of IPLs, each cyclization reduces the number of hydrogen atoms in the elemental composition of alkyl chain by two, resulting in a more oxidized lipid (compare Z_C of chains in Figure 2.2d and e). In this study, the number of internal rings in the GDGT alkyl chains increased downstream in all hot springs, except for Empress Pool, as shown in Figure 2.8. This trend agrees with previous observations of GDGTs sampled spatially along the Bison Pool outflow channel (Schubotz *et al.*, 2013). However, it does not corroborate with the study by Schouten *et al.* (2002) which proposed a positive correlation between GDGT chain cyclization and temperature as the basis for the TEX_{86} paleothermometer, (though TEX_{86} was originally calibrated for sea surface temperatures). Additionally, these observations do not match the results of laboratory growth experiments demonstrating increased chain cyclization with temperature in thermoacidophilic archaea (Boyd *et al.*, 2011). Conflicting trends have also been reported in terrestrial hydrothermal systems. Kaur *et al.* (2015) observed an increase in GDGT ring abundance with temperature between a pH range of 5.5 - 7.2 in hot springs from the Taupo volcanic zone in New Zealand, though their low pH samples did not fit this trend with temperature. Wu *et al.* (2013) studied GDGTs in Yunnan hot springs, China, and found that ring index increased with temperature in one statistical grouping and increased with acidity in the other. Temperature and pH have been cited as competing variables in numerous studies of GDGT ring index in hydrothermal systems and thermophilic archaea (Boyd *et al.*, 2013; Pearson *et al.*, 2008; Boyd *et al.*, 2011). It is still unclear how combinations of geochemical variables influence the incorporation of rings into archaeal membranes in a predictable way; I propose that redox state may influence ring distributions (see the Discussion).

Alkyl chains bearing a secondary hydroxyl group (*e.g.*, Figure 2.2i) were found to

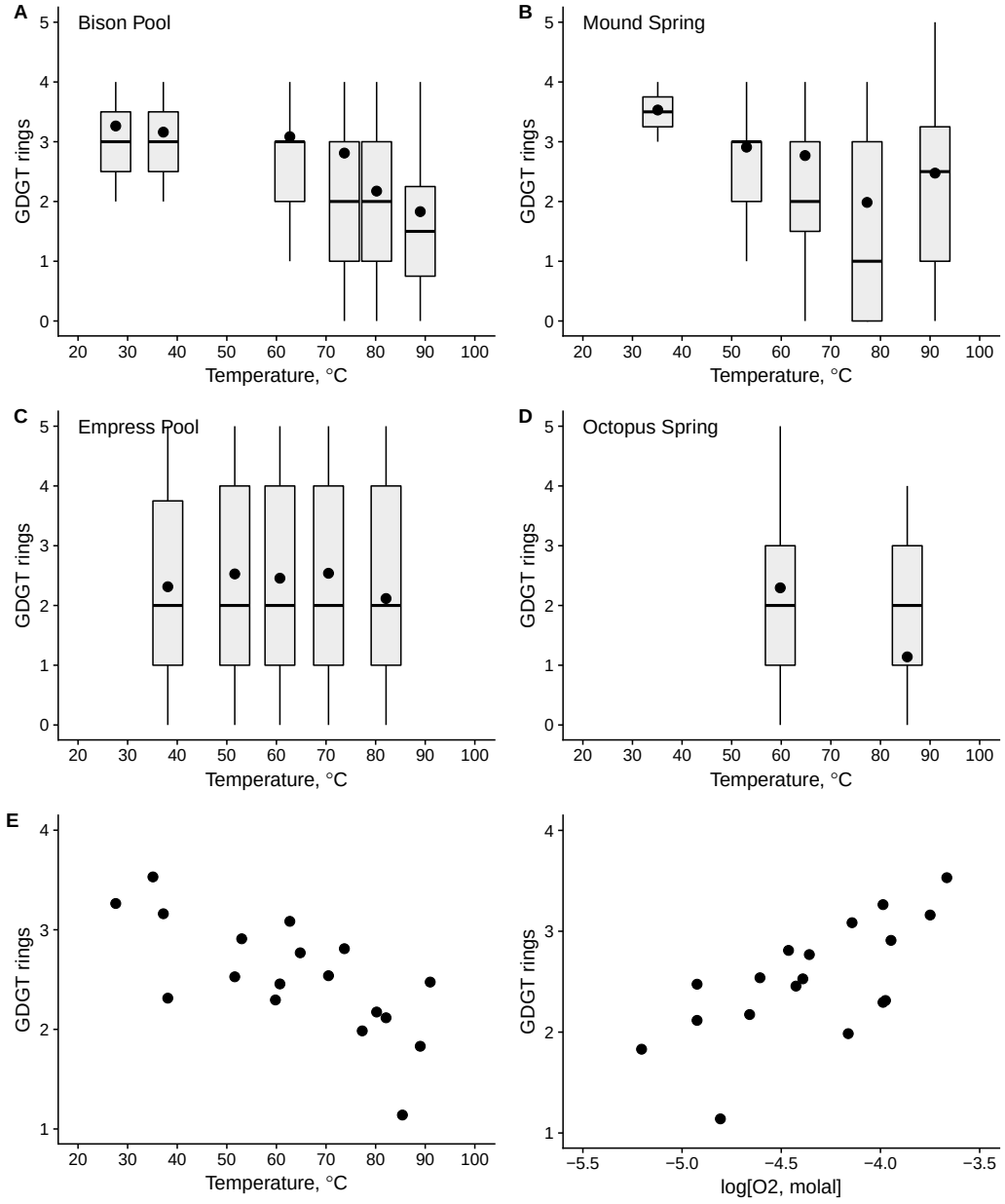


Figure 2.8: Number of internal rings per GDGT. Refer to the caption of Figure 2.6 for a description of the symbols and statistics used in panels A-D. The abundance-weighted number of internal rings per GDGT is shown for all samples in panel E as a function of temperature (left) and dissolved oxygen concentration (right).

comprise a small portion (< 1%) of total chains, as shown by the value of nOH in Table 2.5. Even if the backbone hydroxylation of CER (Figure 2.1d) and FAHFAm-OH (Figure 2.1f) are counted as chain hydroxylations, the total proportion would rise to at most 4% of chains in any given sample. As such, nOH did not substantially affect the weighted Z_C for IPLs in any sample. However, it is possible that alkyl chain hydroxylations are underrepresented. Lipo(oligo/poly)saccharides (LOS/LPS), also known as endotoxins, are lipids rich in fatty acid chains with secondary hydroxylations in the 3' carbon position. LSPs are commonly produced by gram-negative bacteria and may compose up to 75% of their outer membrane (Silipo *et al.*, 2010). LOS has also been traced to certain thermophilic bacteria (Di Lorenzo *et al.*, 2014). These IPLs have masses exceeding 2000 Da, and as such, cannot be quantified by the HPLC-MS method employed in this study.

2.5 Discussion

Lipids were most reduced under the hot, reducing conditions representative of samples closest to the spring sources. Lipids and their alkyl chains became progressively more oxidized toward downstream sites as the surrounding water cooled and became increasingly oxidized. Lipid structures in thermophiles are commonly discussed in terms of how they provide stable membranes at high-temperature (Daniel and Cowan, 2000), but it should not be overlooked that these lipid modifications are the result of adaptation to the system. As such, observed distributions of lipids along hot spring outflow channels represent adaptation to temperature, water chemistry, and almost certainly a host of other variables unaccounted for in this work.

If lipid chain modifications with reduced carbon are an adaptation to more reduced conditions, and oxidized modifications are an adaptation to more oxidized conditions, this might explain why similar lipid distributions have been reported along both

thermal and isothermal redox gradients. Previous studies of IPL distributions in water columns and sediments of the Black Sea report many of the same types of changes in alkyl chain structure with depth along a chemocline (increasingly reduced conditions) as were seen when going from oxidized conditions downstream to reduced conditions upstream samples in hot springs; among them, a pronounced shift in ester- to ether-bonding with increasingly deep, reduced conditions (Schröder, 2015; Schubotz *et al.*, 2009).

Schröder (2015) showed that among non-isoprenoidal IPLs, diester lipids 1G and 2G-DAGs comprised the largest proportion of lipids in the shallowest, most oxidized sample, much like what was seen in the downstream photosynthetic samples of hot springs of this work. Ester-dominated lipids gave way to higher abundances of exclusively ether-linked DEG-lipids with depth and increasingly reduced conditions. Ether-linked isoprenoidal lipids, such as GDGTs and archaeols derived from archaea, were abundant in the deepest samples. Again, these observations parallel changes observed in lipids from downstream to upstream along hot spring outflow channels. On average, lipids in the Black Sea water column were reported to have approximately 3-4 unsaturations in oxic and suboxic water samples, approximately 0.5-2 unsaturations in anoxic water, and less than one unsaturation per lipid in sediments. This indicates that, as was observed in this work, unsaturations become more abundant under increasingly oxidized conditions. While the authors did not see a extensive change in aliphatic carbon content in alkyl chains of DEG lipids with depth, they reported a marked increase in carbon content of BL fatty acids. These fatty acids had an average of about 31 aliphatic carbons summed across all chains (about 15.5 carbons per chain) in the two shallowest, most oxic samples where BL was at its most abundant in the water column, and abruptly increased to an average of about 34-35 carbons (about 17 to 17.5 carbons per chain) across the transition to anoxia.

Considering that the deepest samples were rich in GDGTs with 80 carbons summed across all chains (20 carbons per chain according to the alkyl chain division scheme of this work), the transition from deep and reduced to shallow and oxidized lipids appears to be accompanied by a decrease in alkyl chain length similar to that observed downstream in hot spring samples.

A qualitative guess would suggest abundance-weighted Z_C of IPLs was highest in the oxic zone of the Black Sea water column, and decreases gradually, or becomes more reduced, with depth and increasingly reduced conditions. Oxidation potential is not the only variable to which organisms must adapt in the Black Sea water column, and there are certainly others (pressure, salinity, light availability, *etc.*). Membrane adaptation to the system has manifested in parallel with those adaptations found in hydrothermal gradients, undoubtedly because certain alkyl chain modifications function in both systems. Several questions then arise; what drove natural selection to converge upon similar distributions of alkyl chain structures in both systems? What benefit is there to having reduced lipids under reduced conditions and oxidized lipids under oxidized conditions?

A study of microbial communities sampled at Bison Pool found that the Z_C of genome-encoded proteins increased downstream (Dick and Shock, 2011). Carbon in protein sequences was observed to become more oxidized under oxidized conditions as proportionally more oxidized amino acids were incorporated into protein structures. When proteins were grouped by functional class (*i.e.*, hydrolases, permeases, ATPases, *etc.*), the downstream increase in Z_C was approximately parallel between groups. These encoded proteins have evolved to function in the set of temperature and chemical conditions presented by hydrothermal gradient, and like lipids, show this downstream reduced-to-oxidized directionality in the structural adaptations that provide this function. The fitness benefit of having oxidized and reduced proteins under

oxidized and reduced conditions, respectively, was explored in a follow-up study by Dick and Shock (2013). They predicted through thermodynamic analysis that adaptations observed in the amino acid content of proteins at Bison Pool were energetically favorable in the specific temperature and chemical context of their surroundings; that incorporating more reduced amino acids was most cost-effective under reduced conditions and *vice-versa* for oxidized amino acids.

If the Z_C of proteins indicates an energetic benefit, perhaps a thermodynamic analysis would show the same for lipids. This concept serves as the basis for the study described in Chapter 3. In addition to thermodynamic analyses, future work could be comparative studies of lipid structural adaptations found in various thermal, chemical, and redox conditions. These studies could focus on the influence of redox gradients on the distribution of lipid modifications thought to have similar effects on the biophysical properties of lipid membranes.

For instance, a *cis*-unsaturation introduces a bend in an alkyl chain that is thought to increase membrane fluidity by decreasing lipid packing. Incorporation of a cyclopropane ring also introduces a bend in an alkyl chain with similar effect (Zhang and Rock, 2008), but this modification is more reduced than an unsaturation. Methylation at the anteiso-position of alkyl chains may provide an additional strategy for increasing membrane fluidity with a reduced lipid structural modification. A study by (Zhu *et al.*, 2005) showed *Listeria monocytogenes* expresses anteiso fatty acids in response to cold temperature, presumably also to increase membrane fluidity. All of these adaptations are thought to provide a similar function in regulating membrane fluidity, but a *cis*-unsaturation is more oxidized, and therefore perhaps most energetically cost-effective under oxidized conditions, than either a cyclopentane ring or anteiso branching.

Sterols and hopanoids have relatively reduced cycloalkane hydrocarbon structures.

These have been shown to decrease permeability by condensing membranes while simultaneously imparting fluidity under a variety of temperature and chemical conditions (Belin *et al.*, 2018). Studies of gram-negative bacteria have shown that the incorporation of alkyl chains with a *trans*-unsaturation has a condensing effect on membranes that decreases permeability to solutes (Halverson and Firestone, 2000) and improves membrane stability at higher growth temperatures (Heipieper *et al.*, 1996). Once again, these lipids could represent reduced and oxidized strategies for providing similar membrane properties.

Another lipid modification that has been posited to decrease permeability and increase fluidity is the incorporation of rings into the alkyl chains of GDGTs (Sollich *et al.*, 2017). Details about the environmental controls on GDGT rings are still coming to light, and future studies could focus on water chemistry, including redox states, to explain why GDGT rings correlate positively or negatively with pH and/or temperature in some natural systems but not others (Jia *et al.*, 2014; Boyd *et al.*, 2013). This approach could offer a different take on the interpretation of various GDGT ring indices in paleothermometry. Redox considerations could be extended to the study of H-shaped modifications to GDGTs, which, like GDGT rings, are oxidized relative to the non-H structure but may impart a different effect on the biophysical properties of membranes.

Archaeal GDGTs were found in greater abundance near the high temperature, reduced hot spring source of each sampled spring and represented some of the most reduced lipid structures in this study owing primarily to their relatively high nC and exclusively ether-linked chains. Monolayer-forming tetraethers are not exclusive to archaea; however, as a study by Weijers *et al.* (2006) traced abundances of non-isoprenoidal GDGTs to anaerobic bacteria living in peat. It is intriguing that reduced lipid modifications thought to be exclusive to archaea were also adopted by bacte-

ria thriving under low-oxygen conditions, especially since GDGTs were among the most reduced lipids observed in this study and may be energetically cost-effective to produce in anaerobic environments.

These examples of putative functional homology among oxidized and reduced lipid modifications serve as tantalizing targets of future study. I suggest such studies should examine variables that contribute significantly to the chemical composition and redox state of the system so that general trends in observed lipid distributions can be explored in terms of both function and energetic cost.

2.6 Conclusion

The abundance-weighted average oxidation state of carbon in IPLs increases, *i.e.*, becomes more oxidized, with decreased temperature and increased concentration of dissolved oxygen along the outflow channels of four hot springs in YNP. The incorporation of proportionally more oxidized carbon into the structures of IPLs downstream was primarily attributed to observed changes in the distributions of alkyl chain structures, such as decreasing chain length, increasing degree of unsaturation, glycerol dialkyl glycerol tetraether (GDGT) chain cyclization, and to switch from ether- to ester-bonded alkyl chain-backbone linkages. With the possible exception of GDGT rings (the function of which is still being explored; Sollich *et al.*, 2017), these downstream changes in alkyl chain distribution might be interpreted as merely adaptations to provide thermally tolerant membranes, though I argue that these modifications also represent cost-effective solutions to providing membrane function under the chemical conditions experienced by the microbial communities sampled in this study. The tendency to find oxidized lipid modifications under oxidized conditions and *vice versa* for reduced conditions may represent an evolutionary drive to maximize lipid function while minimizing energetic cost.

Chapter 3

THERMODYNAMIC FAVORABILITY OF THERMOPHILE LIPID CHAIN MODIFICATIONS ACROSS A TEMPERATURE AND REDOX GRADIENT

3.1 Introduction

Ongoing efforts are needed to elucidate factors that control lipid distribution in the environment. Our level of understanding ranges widely with respect to how lipid structural adaptations contribute to the biophysical properties or function of the membrane, and the roles played by many lipid structures are still unknown. Lipids need to function effectively for an organism, and it is plausible that lipid distributions offering the best fitness advantage for the lowest energy cost are selected over the course of evolution.

In this study, I performed a thermodynamic investigation to address the hypothesis that organisms adapt their lipid compositions to maximize function while also minimizing energetic cost (see Chapter 1). Specifically, I wanted to explore whether the downstream increase the average oxidation state of carbon (Z_C) in alkyl chains reported in Chapter 2 signals an underlying energetic adaptation in microbial membranes, and hypothesized that **organisms adapt their alkyl chains to maximize function while also minimizing energetic cost**. This study makes use of the same sample set described in Chapter 2, which includes eighteen samples taken across the temperature and chemical gradients of four alkaline hot springs in Yellowstone National Park (YNP); Bison Pool (BP), Mound Spring (MS), Empress Pool (EP), and Octopus Spring (OS). Previous thermodynamic studies on the energetic favorability of proteins along the temperature and redox gradient of Bison Pool, YNP (Dick and

Shock, 2011, 2013) served as a conceptual foundation for the use of equilibrium assumptions for predicting relative stabilities of biomolecules based on temperature and a set of known geochemical variables. These calculations are based on predicted favorability of reactions to form the biomolecules of interest from 'basis species', which in this work were a set of naturally available inorganic aqueous species relevant to autotrophic metabolism and included bicarbonate (HCO_3^-), ammonium (NH_4^+), protons (H^+), electrons (e^-) and water (H_2O) for all but one sample that used nitrate (NO_3^-) in place of NH_4^+.

If a biomolecule is predicted to be more stable relative to biomolecules with homologous function, this might indicate that the biological energy requirement for assembling this biomolecule out of available materials (basis species) is the most cost-effective evolutionary option. Likewise, the stability of entire biomolecular assemblages can be assessed so long as thermodynamic properties for individual compounds exist or can be estimated. I wanted to explore whether observed distributions of intact polar lipid (IPL) alkyl chains represented stable, cost-effective assemblages in the context of hot spring temperature and chemical composition. I reasoned that if the alkyl chains sampled from high-temperature, reduced conditions were predicted to be more stable under those conditions relative to the alkyl chains sampled from other sites, and *vice versa* for oxidized conditions, then this would provide evidence of energetically driven convergence on functional lipids.

The thermodynamic properties of alkyl chains had to be obtained first, but the scarcity of experimentally derived literature data presented a challenge. Estimation methods exist for deriving properties of organic molecules assuming the chemical structures are known, but the structures of individual alkyl chains cannot be resolved with the analytical method used to quantify IPLs in Chapter 2. This inspired the use of 'sample-averaged free alkyl chains' ($alkyl_{free}$) as a model for observed lipid

distributions in a sample.

Various abundance-weighted structural characteristics of alkyl chains such as nC (number of aliphatic carbons), nUnsat (number of unsaturations), nOH (number of secondary hydroxylations), mole fraction of monolayer-forming chains and backbone-chain linkage types (ether, ester, amide, and C-C) were calculated from observed abundances of IPLs reported in Table 2.5 of Chapter 2. In this study, partial molal thermodynamic aqueous properties of eighteen hypothetical 'sample-averaged' alkyl chains were estimated, each representing the abundance-weighted structural characteristics of alkyl chains in a hot spring sample. For instance, the nC and nUnsat in the alkyl chains of Bison Pool sample BP1 were observed to be 19.80 and 3.12×10^{-1}, respectively, so the thermodynamic properties estimated for the aqueous sample-averaged alkyl chain representing BP1 assume a hypothetical chemical structure with 19.80 aliphatic carbons and 3.12×10^{-1} unsaturations, and so on. The chemical cartoon shown in Figure 3.1 represents one possible visualization of the hypothetical structure of the sample-averaged alkyl chain of sample BP1. Calculating the properties of this sample-averaged molecule is equivalent to calculating thermodynamic properties for thousands of individual chains and then taking an abundance-weighted average of each property. It would not have been possible to estimate the thermodynamic properties of every observed chain in this study due to an inability to resolve structural characteristics of individual IPL alkyl chains using this LC-MS method. The rationale for calculating thermodynamic properties from abundance-weighted structural characteristics was as follows: if an IPL observed in a sample has 38 aliphatic carbons and two unsaturations distributed between two chains, the structural characteristics of its chains must be described using averages without resorting to assumptions about which chain bears which combination of characteristics, and by extension, the estimated thermodynamic properties of alkyl chains must therefore be described by

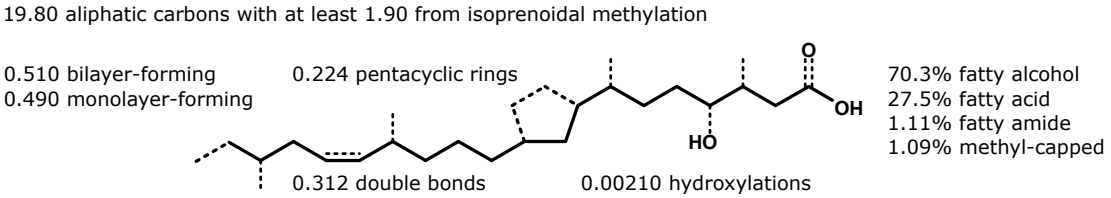


Figure 3.1: Hypothetical chemical structure of the sample-averaged free alkyl chain ($\text{alkyl}_{\text{free}}$) of Bison Pool sample BP1 based on abundance-weighted structural characteristics of alkyl chains linked to observed intact polar lipids. Position and isomerism of hydroxylations, pentacyclic rings, methyl branches, and double bonds are not considered when calculating thermodynamic properties of $\text{alkyl}_{\text{free}}$; their depiction here is to illustrate one of many possible interpretations of a hypothetical average molecule.

averages as well. This was taken a step further by assuming the structural characteristics of all alkyl chains in a sample could be averaged for the sake of calculating thermodynamic properties, resulting in the concept of a 'sample-averaged' alkyl chain.

In this work, thermodynamic properties were estimated for 'free' alkyl chains, as opposed to IPL-linked chains, primarily to provide a simpler model for representing lipid distribution. The difference between IPL-linked and free alkyl chains is represented in Figure 3.2 for ether-, ester-, amide-, and C-C-linked chains. Except for the C-C linkage type, free chains represent a hydrolyzed version of the IPL-linked chains, with fatty acids, alcohols, and amides corresponding to the free versions of IPL-linked esters, ethers, and amide-bonded alkyl chains. In the case of C-C-linked alkyl chains, the free chain was methyl-capped to maintain its alkane-like properties.

A study by Shock and Helgeson (1990) demonstrated that the partial molal thermodynamic properties of aqueous hydrocarbons tend to have a strong linear correlation with chain length, as was shown for 1-alkynes, n-alkyl benzenes, 1-alkenes, 1-amines, n-alkanes, aldehydes, 1-alcohols, 2-ketones, methyl alkanoates, n-carboxylic acids, and n-carboxylate ions. This observation served as a theoretical foundation for estimating properties of aqueous $\text{alkyl}_{\text{free}}$ with non-integer values for nC in this work. If a thermodynamic property is linearly correlated with chain length, it should make no difference whether a chain length of 19.0 or 19.8 is used to solve for the value of the property.

Shock and Helgeson (1990) also showed thermodynamic properties regressed against chain length were approximately parallel for different functional group chains and differed only by y-intercept. Because the slopes of these regressions were nearly identical, each additional -CH_2- group was suggested to contribute the same change in the value of the property. This strongly suggests an additive nature for both functional groups and -CH_2- in linear organic molecules.

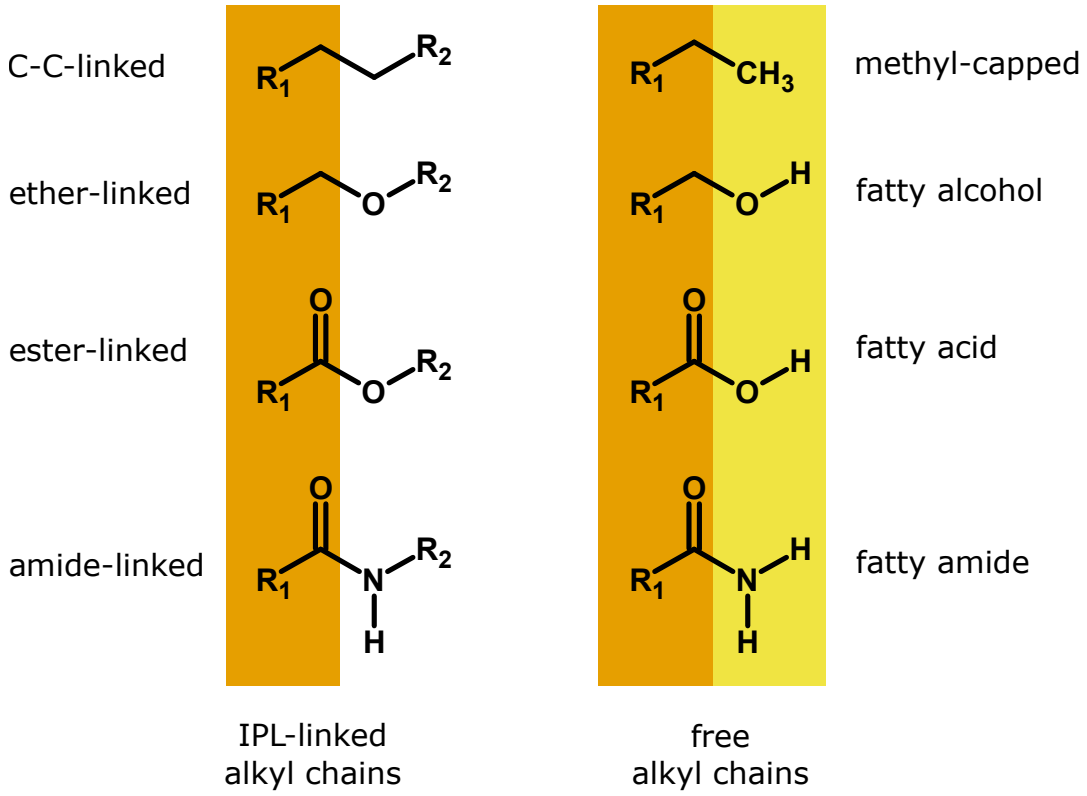


Figure 3.2: Comparison of IPL backbone-alkyl chain linkage types (left) and the free alkyl chains used in the thermodynamic calculations of this study (right). R_1 indicates the rest of the alkyl chain and R_2 stands for the remainder of the backbone, headgroup, and other alkyl chains of the IPL. As a part of the thermodynamic estimations performed in this study, abundance-weighted chemical formulae and structures of IPL-linked alkyl chains (orange box, given in Chapter 2) were capped with the elemental and thermodynamic contribution of groups in the yellow box according to alkyl chain-backbone linkage type, arriving at 'free' sample-averaged alkyl chains as a model for sample lipid distributions.

This theory of adding together the contribution of individual chemical groups to arrive at the estimated properties of the entire molecule, or 'group additivity', is not new. Sugden (1924) showed a property related to surface tension, parachor, could be estimated by group additivity for a variety of organic and inorganic compounds, as well as for a few elements. Benson and Buss (1958) noted that various group additivity rules had cropped up in the literature and proposed a series generalized of rules for estimating thermodynamic properties of ideal gases, including heat capacity, entropy, and heat of formation from the elements. Studies by Cabani *et al.* (1981, 1977a,b) measured volumetric and thermodynamic properties of aqueous molecules and calculated group contribution values of functional groups. Since then, group additivity has been used to estimate the properties of a great many aqueous organic molecules that lack experimentally measured values.

According to group additivity theory, the thermodynamic contribution of one functional group to a molecule is expected to be roughly the same for each addition. This suggests a linear trend in the estimated property as a function of the number of times the functional group is added, just as was described above for CH_2 groups. Following the same reasoning that was applied to estimating the properties of a chain with a non-integer nC, I estimated the contributions of non-integer unsaturations, hydroxylations, GDGT cyclic rings, isoprenoidal branching to the thermodynamic properties of $\text{alkyl}_{\text{free}}$. The fractional contributions of the carboxylic acid, alcohol, methyl-capped, and amide functional group of $\text{alkyl}_{\text{free}}$ were also estimated using these guidelines.

3.2 Methods

3.2.1 Analysis of water chemistry

Sampling locations were measured and water samples were collected, filtered, and analyzed as described in Chapter 2. Briefly, temperature and pH were measured in the field as close to the sampling locations as possible using a YSI 30 conductivity meter for temperature and a model 3300i or 3110 WTW pH meter with WTW probe for pH. Concentrations of dissolved oxygen and sulfide were measured in unfiltered water samples using Hach 2400 or 2800 portable spectrophotometer with Hach reagents. Water samples collected for later laboratory analyses were filtered with Supor^{TM} (Pall Corporation) polyethersulfone (PES) filters down to 0.2 μm; samples for ion chromatography were collected in 30 mL HDPE Nalgene bottles (two per sample) and stored at -20°C before analysis on Dionex DX-600 ion chromatography systems for concentrations of major cations and anions.

Filtered samples for dissolved inorganic carbon (DIC) analysis were collected in acid-washed amber glass vials with black butyl rubber septa. DIC was chemically converted to CO_2 by reacting with phosphoric acid before analysis on an OI Wet Oxidation TOC analyzer coupled to a Thermo Delta Plus Advantage mass spectrometer as described in Havig *et al.* (2011).

The reduction potential (Eh) of water samples was calculated using measured concentrations of major dissolved solutes and hot spring temperatures using EQ3/6 aqueous geochemical speciation software version 8.0a (Wolery, 2002) and thermodynamic data from the slop16 database available for download at the URL http://geopig.asu.edu/sites/default/files/slop16.dat.

3.2.2 Analysis of environmental IPLs

Methods used to collect, identify, and quantify the hot spring microbial IPLs from sediments and biofilms are described in detail in Chapter 2. Briefly, sediments and biofilms were collected into sterile specimen containers with sterilized forceps and spatulas, placed on dry ice the field, and stored at -80°C after transport to the lab. Samples were freeze-dried, homogenized with mortar and pestle, lipids were extracted using a modified version of the Bligh and Dyer method (White and Ringelberg, 1998). Aliquots of the resulting total lipid extracts (TLEs) were analyzed using the hydrophilic interaction chromatography (HILIC) method (Wörmer *et al.*, 2013) on an Agilent 1200 series high-performance liquid chromatograph (HPLC) attached to an Agilent 6520 Accurate-Mass Quadrupole Time-of-Flight Mass Spectrometer with an electrospray ionization source operated in positive ion mode. Commercially available standards were used to obtain response factors for calculating IPL mole fractions from mass spectral peak areas. Response factors are given in Table A.1 and peak assignments to observed IPLs is reported in Table 2.1.

3.2.3 Deriving properties and chemical formulae of $alkyl_{free}$

Calculation of the elemental abundance of $alkyl_{free}$ began with the chemical formulae of sample-averaged IPL-linked alkyl chains, to which additional carbon, hydrogen, oxygen, and nitrogen were added to convert ether-linked chains into free fatty alcohols, ester-linked chains into free fatty acids, amide-linked chains into free fatty amides, and C-C-linked chains into free methyl-capped chains. Chemical formulae of sample-averaged IPL-linked alkyl chains and mole fractions of alkyl chain linkage types required for these calculations are reported in Table 3.1 for isoprenoidal chains and in Table 2.5 of Chapter 2 for all others.

Table 3.1: Summary of abundance-weighted structural characteristics of isoprenoidal alkyl chains.

Site	Sample	x_{AR}[a]	x_{GDGT}[b]	nPent[c]	nHex[d]
Bison	BP1	1.98×10^{-2}	4.90×10^{-1}	2.24×10^{-1}	
Pool	BP2	2.30×10^{-2}	3.00×10^{-1}	1.63×10^{-1}	
	BP3	1.49×10^{-3}	8.81×10^{-2}	6.19×10^{-2}	
	BP4		7.16×10^{-3}	5.52×10^{-3}	
	BP5	5.06×10^{-6}	7.20×10^{-4}	5.69×10^{-4}	
	BP6		2.94×10^{-3}	2.40×10^{-3}	
Mound	MS1		9.99×10^{-1}	6.16×10^{-1}	2.68×10^{-3}
Spring	MS2	2.30×10^{-2}	6.07×10^{-1}	3.01×10^{-1}	
	MS3		1.64×10^{-1}	1.13×10^{-1}	
	MS4		5.85×10^{-2}	4.25×10^{-2}	
	MS5		1.09×10^{-2}	9.65×10^{-3}	
Empress	EP1	1.97×10^{-2}	8.63×10^{-1}	4.45×10^{-1}	1.09×10^{-2}
Pool	EP2	4.74×10^{-3}	4.71×10^{-1}	2.91×10^{-1}	7.83×10^{-3}
	EP3	2.21×10^{-3}	6.06×10^{-1}	3.61×10^{-1}	1.15×10^{-2}
	EP4	4.08×10^{-3}	2.31×10^{-1}	1.43×10^{-1}	2.98×10^{-3}
	EP5	2.78×10^{-3}	1.26×10^{-1}	7.19×10^{-2}	1.12×10^{-3}
Octopus	OS1	1.26×10^{-1}	4.21×10^{-1}	1.20×10^{-1}	
Spring	OS2	5.08×10^{-6}	7.60×10^{-3}	4.21×10^{-3}	1.46×10^{-4}

Note: a blank value indicates the structural characteristic was not detected in alkyl chains in the sample.
[a] Mole fraction of alkyl chains belonging to archaeol, [b] mole fraction of alkyl chains belonging to glycerol dialkyl glycerol tetraethers, [c] weighted average number of cyclopentane rings per alkyl chain, [d] weighted average number of cyclohexane rings per alkyl chain.

The number of carbon atoms (c_{free}) in the chemical formulae of alkyl$_{\text{free}}$ was calculated with:

$$c_{free} = c_{linked} + x_{cc}, \tag{3.1}$$

where c_{linked} indicates the number of carbon atoms in the sample-averaged IPL-linked alkyl chain of the sample, and x_{cc} stands for the observed mole fraction of C-C-linked alkyl chains. Addition of a number of carbons equal to x_{cc} to c_{linked} accounts for the carbon that would be gained with a terminal methyl group, assuming hypothetical carbon-carbon bond breakage of C-C-linked alkyl chains to produce the sample-averaged free chains considered in this study.

The number of hydrogen atoms (h_{free}) in the chemical formulae of alkyl$_{\text{free}}$ was calculated from

$$h_{free} = h_{linked} + x_{ether} + x_{ester} + 2(x_{amide}) + 3(x_{cc}), \tag{3.2}$$

where h_{linked} represents the number of hydrogen atoms in the sample-averaged IPL-linked alkyl chain of the sample, and x_{ether}, x_{ester}, and x_{amide} correspond to the observed mole fractions of ether-, ester-, and amide-linked alkyl chains, respectively. A number of hydrogens equal to x_{ether} and x_{ester} are added to h_{linked} to account for the single hydrogen in the hydroxyl groups of the resulting alcohol and carboxylic acid group assuming hypothetical hydrolysis of ether- and ester-linked chains to produce free fatty acids and fatty alcohols. Similarly, two times x_{amide} is added to account for the two hydrogen atoms in the amine group of the free amide after the hypothetical hydrolysis of an amide-linked chain. Lastly, adding a number of hydrogens equal to three times x_{cc} accounts for the three hydrogen atoms added by a terminal methyl group after the hypothetical carbon-carbon breakage of a C-C-linked alkyl chain.

The number of nitrogen atoms (n_{free}) in the chemical formulae of $\text{alkyl}_{\text{free}}$ was calculated as

$$n_{free} = n_{linked} + x_{amide}, \tag{3.3}$$

where n_{linked} indicates the number of nitrogen atoms in the sample-averaged IPL-linked alkyl chain. A number of nitrogen atoms equal to x_{amide} are added to n_{linked} to account for the single nitrogen atom added by the amine group of the free amide after hypothetical hydrolysis of an amide-linked alkyl chain.

The number of oxygen atoms (o_{free}) in the chemical formulae of $\text{alkyl}_{\text{free}}$ was calculated with

$$o_{free} = o_{linked} + x_{ether} + x_{ester}, \tag{3.4}$$

where o_{linked} is the number of oxygen atoms in the sample-averaged IPL-linked alkyl chain. A number of oxygen atoms equal to x_{ether} and x_{ester} are added to o_{linked} to account for the single oxygen atom in the hydroxyl groups of the resulting alcohol and carboxylic acid group assuming hypothetical hydrolysis of ether- and ester-linked chains to produce free fatty acids and fatty alcohols.

3.2.4 Calculation of thermodynamic properties

Thermodynamic properties for $\text{alkyl}_{\text{free}}$ were estimated assuming an aqueous phase. The standard state convention for species in the aqueous phase assumed unit activity in a hypothetical 1 mol$\times$kg^{-1} solution referenced to infinite dilution at any temperature and pressure. The standard state used for liquid water was one unit activity of pure H_2O at any temperature and pressure. The standard state used for gaseous species was unit fugacity of the ideal gas at 1 bar and any temperature.

Hydration properties were chosen to reflect the equilibrium process of transferring 1 mole of the species in the gas phase to the aqueous phase (Plyasunov and Shock, 2000).

Aqueous partial molal thermodynamic Gibbs free energies and enthalpies of formation from the elements ($\Delta_f G^\circ_{aq}$ and $\Delta_f H^\circ_{aq}$, respectively), volumes (V°_{aq}), heat capacities (Cp°_{aq}), and Gibbs free energies of the hydration process ($\Delta_h G^\circ$) were estimated for $\text{alkyl}_{\text{free}}$ in two steps. The goal of the first step was to estimate the properties of a saturated, straight aliphatic chain of length nC, terminating at one end with a hybrid functional group. This hybrid group was a combination of alcohol, carboxylic acid, alkane, and amide, with the fraction of each set by observed mole fractions of ether, ester, C-C, and amide-linked IPL alkyl chains, respectively.

Experimentally derived thermodynamic properties of carboxylic acids, primary alcohols, and alkanes were gathered from the literature into Table 3.2 and then regressed as a linear function of chain length. The results of these regressions are shown in Table 3.3. Properties for n-amides were not readily available and were estimated with a different method, described later. The slopes and intercepts of these regressions were used to estimate the properties of a hybrid fatty alkyl chain, Ξ_{fatty}, according to the equation:

$$\begin{aligned}\Xi_{fatty} =&(m_{alcohol} \cdot \text{nC} + b_{alcohol})(x_{ether}) \\ &+(m_{acid} \cdot \text{nC} + b_{acid})(x_{ester}) \\ &+(m_{alkane} \cdot \text{nC} + b_{alkane})(x_{cc}) \\ &+(m_{acid} \cdot \text{nC} + b_{acid} + \Delta\Xi_{amide})(x_{amide}),\end{aligned} \tag{3.5}$$

where $m_{alcohol}$, m_{acid}, m_{alkane}, $b_{alcohol}$, b_{acid}, and b_{alkane} correspond to the slopes (m) and intercepts (b) of linear regressions of the value of each thermodynamic property as a function of chain length for fatty acids, fatty alcohols, and alkanes. Mole fractions

Table 3.2: Partial molal thermodynamic aqueous and hydration properties of aliphatic carboxylic acids, alcohols, and alkanes with nC > 2 used in the estimation of saturated fatty acid, fatty alcohol, and fatty hydrocarbon properties.

	$\Delta_f G^{\circ}_{aq}$ [a]	$\Delta_f H^{\circ}_{aq}$ [a]	Cp°_{aq} [b]	V°_{aq} [c]	$\Delta_h G^{\circ}$ [a]
carboxylic acids[d]					
propanoic acid	-390.99	-512.41	253	67.9	-21.4[e]
n-butanoic acid	-381.62	-535.34	337	84.61	-20.9[e]
n-pentanoic acid	-373.38	-559.36	432.2	100.5	-19.0[e]
n-hexanoic acid	-364.51	-582.79	523.0	116.55	-17.8[e]
n-heptanoic acid	-356.48	-607.01	611.7	131.6	-[f]
n-octanoic acid	-349.15	-631.99	700.4	147.4	-
n-nonanoic acid	-339.82	-654.92	789.1	163.2	-
n-decanoic acid	-331.25	-678.64	877.8	179.0	-
n-undecanoic acid	-322.71	-702.37	966.5	194.8	-
n-dodecanoic acid	-314.13	-726.09	1055	210.6	-
primary alcohols[g]					
1-propanol	-172.07	-312.74	355	70.71	-12.36
1-butanol	-162.40	-336.53	443	86.53	-11.86
1-pentanol	-151.72	-359.56	535	102.50	-11.11
1-hexanol	-143.31	-383.79	617	118.46	-10.30
1-heptanol	-134.41	-408.39	702	133.03	-9.72
1-octanol	-123.93	-429.85	778	151.20	-9.14
1-nonanol	-115.24	-454.67	865	164.55	-8.36
1-decanol	-106.23	-478.75	950	180.16	-8.18
1-undecanol	-96.52	-502.77	1034	195.77	-6.90
1-dodecanol	-86.17	-525.73	1119	211.38	-6.04
alkanes[g]					
propane	-8.22	-127.43	398	69.31	16.09
n-butane	0.24	-151.53	482	75.55	16.58
n-pentane	8.88	-175.45	573	99.07	17.52
n-hexane	19.26	-198.69	636	114.68	18.26
n-heptane	27.45	-224.54	739	130.29	18.95
n-octane	36.08	-250.45	824	145.90	19.64
n-nonane	46.24	-275.24	908	161.51	20.51
n-decane	55.50	-296.34	993	177.12	21.24
n-undecane	64.71	-321.12	1078	192.73	22.94
n-dodecane	74.65	-343.26	1163	208.34	22.51
n-tridecane	81.58	-368.84	1248	223.95	22.87
n-tetradecane	90.85	-392.99	1333	239.56	23.55

[a] kJ mol^{-1}, [b] J mol^{-1} K^{-1}, [c] cm^3 mol^{-1}, [d] carboxylic acid properties are taken from Shock (1995) unless otherwise noted, [e] values calculated from Henry's law constants, K_H, reported in Khan and Brimblecombe (1992) using the relation $\Delta_h G^{\circ} = -RT \ln K_H$, [f] indicates a value that was not found in the literature, [g] properties for alcohols and alkanes are recommended values from the ORganic Compounds HYDration properties (ORCHYD) database compiled by Plyasunova *et al.* (2004).

Table 3.3: Regression statistics for partial molal properties of fatty acids, fatty alcohols, and alkanes with nC > 2 as a function of carbon length.

property	chain type	slope	intercept	R^2
$\Delta_f G^{\circ}_{aq}$ [a]	fatty acid	8.46	-415.87	>0.99
	fatty alcohol	9.43	-199.95	>0.99
	alkane	9.10	-35.95	>0.99
$\Delta_f H^{\circ}_{aq}$ [a]	fatty acid	-23.82	-440.45	>0.99
	fatty alcohol	-23.70	-241.52	>0.99
	alkane	-24.14	-55.30	>0.99
Cp°_{aq} [b]	fatty acid	89	-15	>0.99
	fatty alcohol	84	108	>0.99
	alkane	85	140	>0.99
V°_{aq} [c]	fatty acid	15.8	21.3	>0.99
	fatty alcohol	15.61	24.36	>0.99
	alkane	15.80	18.84	>0.99
$\Delta_h G^{\circ}$ [a]	fatty acid	1.3	-25.5	0.96
	fatty alcohol	0.68	-14.52	0.99
	alkane	0.72	13.97	0.98

Note: regressed thermodynamic properties are given in Table 3.2. F-test p-values were significant for all linear models. [a] kJ mol^{-1}, [b] J mol^{-1} K^{-1}, [c] cm^3 mol^{-1}

of ether-linked, ester-linked, C-C-linked, and amide-linked alkyl chains observed in the sample are indicated by x_{ether}, x_{ester}, x_{cc}, and x_{amide}, respectively. The estimated difference in the thermodynamic property between a fatty amide and fatty acid is represented by $\Delta\Xi_{amide}$, given in Table 3.4. The first three terms of this equation use the slopes and intercepts of the regression of the property of interest to obtain the property of a straight fatty chain with a length equal to nC, and then weights the contributions of those properties by the mole fractions of observed IPL alkyl chain linkages; fatty alcohols, fatty acids, and fatty alkanes are weighted by the mole fractions of ethers, esters, and C-C linked alkyl chains, respectively. In the fourth term of the equation, a carboxylic acid of length nC was instead estimated by $m_{acid} \cdot \mathrm{nC} + b_{acid}$, to which $\Delta\Xi_{amide}$ was added to convert the carboxylic acid into an amide. The contribution of fatty amides to Ξ_{fatty} was not estimated by regression of aqueous amide properties in the same way as alcohols, carboxylic acids, and alkanes due to a lack of literature data, necessitating the carboxylic acid to amide correction, *i.e.*, $\Delta\Xi_{amide}$.

The thermodynamic contributions of unsaturation, chain hydroxylation, pentacyclic ring, hexacyclic ring, monolayer-forming, and isoprenoidal methylation properties were subsequently added to Ξ_{fatty}, each weighted by their observed per-chain abundance with the equation:

Table 3.4: Partial molal thermodynamic aqueous and hydration properties of functional groups and chemical species used in the estimation of fatty amide properties.

	$\Delta_f G^{\circ}_{aq}$ [a]	$\Delta_f H^{\circ}_{aq}$ [a]	Cp°_{aq} [b]	V°_{aq} [c]	$\Delta_h G^{\circ}$ [a]
[-$CONH_2$]	-194.8[d]	-257.0[d]	21[d]	28.2[d]	
[-COOH]	-387.3[d]	-436.7[d]	22[d]	24.5[d]	
acetamide					-32.7[e]
acetic acid					-21.0[f]
$\Delta\Xi_{amide}$	192.5[g]	179.7[g]	-1[g]	3.7[g]	-11.7[h]

[a] kJ mol^{-1}, [b] J mol^{-1} K^{-1}, [c] cm^3 mol^{-1}, [d] Dick *et al.* (2006), [e] value from Cabani *et al.* (1981) with an adjustment of +7.9 kJ/mol to convert the standard state definition used by the authors for gases (1M ideal gas) to those used here (pure ideal gas) when deriving properties of the hydration process, as reported in the *Database on the infinite dilution partial molar Gibbs energies of hydration of aqueous organic nonelectrolytes* (Sedlbauer, 2012; Sedlbauer *et al.*, 2002), [f] values calculated from Henry's law constants, K_H, reported in Khan and Brimblecombe (1992) using the relation $\Delta_h G^{\circ} = -RT \ln K_H$, [g] estimated from $\Xi_{[-CONH_2]} - \Xi_{[-COOH]}$, where $\Xi_{[-CONH_2]}$ and $\Xi_{[-COOH]}$ are the properties of [-$CONH_2$] and [-COOH] given in the table, respectively, [h] estimated from $\Xi_{acetamide} - \Xi_{acetate}$, where $\Xi_{acetamide}$ and $\Xi_{acetate}$ are the properties of acetamide and acetate given in the table, respectively.

$$\begin{aligned}\Xi_{alkyl,free} = \Xi_{fatty} &+(\Delta\Xi_{unsat})(\text{nUnsat}) \\ &+(\Delta\Xi_{hydroxyl})(\text{nOH}) \\ &+(\Delta\Xi_{pent})(\text{nPent}) \\ &+(\Delta\Xi_{hex})(\text{nHex}) \\ &+(\Delta\Xi_{monolayer})(x_{GDGT}) \\ &+(\Delta\Xi_{methyl})(\text{nMethyl}),\end{aligned} \tag{3.6}$$

where $\Xi_{alkyl,free}$ designates the thermodynamic property of the $\text{alkyl}_{\text{free}}$, x_{GDGT} corresponds to the mole fraction of alkyl chains belonging to GDGTs given in Table 3.1, nC, nUnsat, nHydrox, nPent, and nHex stand for the abundance-weighted average number of aliphatic carbons, unsaturations, hydroxylations, pentacyclic rings, and hexacyclic per alkyl chain observed in the sample, respectively, $\Delta\Xi_{unsat}$ designates the estimated difference in the thermodynamic property between a monounsaturated and saturated chain taken from Table 3.5,

$\Delta\Xi_{hydroxyl}$ indicates the estimated difference in the thermodynamic property between a hydroxylated and nonhydroxylated chain taken from Table 3.6, $\Delta\Xi_{pent}$ represents the estimated difference in the thermodynamic property between a chain with a pentacyclic ring and a noncyclic chain taken from Table 3.7, $\Delta\Xi_{hex}$ stands for the estimated difference in the thermodynamic property between a chain with a hexacyclic ring and a noncyclic chain taken from Table 3.7, $\Delta\Xi_{monolayer}$ corresponds to the estimated difference in the thermodynamic property between a monolayer-forming and bilayer-forming chain taken from Table 3.6, $\Delta\Xi_{methyl}$ designates the estimated difference in the thermodynamic property between a methyl-branched and straight chain taken from Table 3.6, and nMethyl represents the abundance-weighted number of methyl branches per alkyl chain. In this study, methyl branching could not be

Table 3.5: Partial molal thermodynamic data for alkenes used in the estimation of unsaturated alkyl chain properties.

	$\Delta_f G^{\circ}_{aq}$ [a]	$\Delta_f H^{\circ}_{aq}$ [a]	Cp°_{aq} [b]	V°_{aq} [c]	$\Delta_h G^{\circ}$ [a]
2-pentene	83.09	-59.2	535	87.00	13.44
2-heptene	99.65	-109.3	709	118.40	14.77
$\Delta\Xi_{unsat}$	73.205[d]	115.7[d]	-34[d]	-11.98[d]	-4.13[d]

Properties in this table are recommended values from the ORganic Compounds HYDration properties (ORCHYD) database compiled by Plyasunova *et al.* (2004) unless otherwise noted. [a] kJ mol^{-1}, [b] J mol^{-1} K^{-1}, [c] cm^3 mol^{-1}, [d] estimated by taking the average of $\Xi_{2-pentene} - \Xi_{pentane}$ and $\Xi_{2-heptene} - \Xi_{heptane}$, where $\Xi_{2-pentene}$, and $\Xi_{2-heptene}$ are the properties of 2-pentane and 2-heptene given in the table and $\Xi_{pentane}$, and $\Xi_{heptane}$ are the properties of pentane and heptane given in Table 3.2.

Table 3.6: Partial molal thermodynamic data for second order groups used in the estimation of alkyl chain methyl branching, hydroxylation, and monolayer-spanning characteristics.

Group	Formula	$\Delta_f G^{\circ}_{aq}$ [a]	$\Delta_f H^{\circ}_{aq}$ [a]	S°_{aq} [b]	Cp°_{aq} [b]	V°_{aq} [c]	$\Delta_h G^{\circ}$ [a]	$\Delta_h H^{\circ}$ [a]	$\Delta_h S^{\circ}$ [b]	$\Delta_h Cp^{\circ}$ [b]	$\Delta_f H^{\circ}_{ig}$ [a]	S°_{ig} [b]	Cp°_{ig} [b]
C-$(H)_2(C)_2$	CH_2	9.05[d]	-24.15[e]	25.07[f]	85[g]	15.61[h]	0.68[h]	-3.52[h]	-14.09[i]	62[h]	-20.63[j]	39.16[j]	22.89[j]
C-$(H)(C)_3$	CH	34.07[d]	1.17[e]	-39.28[f]	3[g]	5.96[h]	-1.93[h]	2.34[h]	14.32[i]	-17[h]	-1.17[j]	-53.60[j]	20.08[j]
C-$(H)_3(C)$	CH_3	-16.35[d]	-50.45[e]	87.37[f]	158[g]	25.56[h]	3.72[h]	-8.19[h]	-39.95[i]	132[h]	-42.26[j]	127.32[j]	25.73[j]
C-$(H)(O)(C)_2$	CH	6.29[d]	-27.98[e]	-43.85[f]	26[g]	7.48[h]	-1.64[h]	-1.88[h]	-0.80[i]	6[h]	-26.10[j]	-43.05[j]	19.96[j]
O-(H)(C)	OH	-170.59[d]	-197.67[e]	78.30[f]	27[g]	11.35[h]	-25.46[h]	-38.34[h]	-43.20[i]	9[h]	-159.33[j]	121.5[j]	18.16[j]
$\Delta\Xi_{methyl}$	$2\times CH_2 \rightarrow CH(CH3)$	0.38[k]	-0.98[k]		9[k]	0.30[k]	0.43[k]						
$\Delta\Xi_{hydroxyl}$	$CH_2 \rightarrow CH(OH)$	-173.35[l]	-201.50[l]		-32[l]	3.22[l]	-27.78[l]						
$\Delta\Xi_{monolayer}$	$CH_3 \rightarrow CH_2$	25.40[m]	26.30[m]		-73[m]	9.65[m]	-3.04[m]						

[a] kJ mol^{-1}, [b] J mol^{-1} K^{-1}, [c] cm^3 mol^{-1}, [d] calculated using $\Delta_f H^{\circ}_{aq}$ and S°_{aq} given in the table along with S° of the elements referenced from Cox *et al.* (1989) and the relation $\Delta_f G^{\circ}_{aq} = \Delta_f H^{\circ}_{aq} - T\Delta(S^{\circ}_{aq} - \sum S^{\circ}_{elements})$, [e] calculated using $\Delta_h H^{\circ}$ and $\Delta_f H^{\circ}_{ig}$ given in the table and the relation $\Delta_f H^{\circ}_{aq} = \Delta_h H^{\circ} + \Delta_f H^{\circ}_{ig}$, [f] calculated using $\Delta_h S^{\circ}$ and S°_{ig} given in the table and the relation $S^{\circ}_{aq} = \Delta_h S^{\circ} + S^{\circ}_{ig}$, [g] calculated using $\Delta_h Cp^{\circ}$ and Cp°_{ig} given in the table and the relation $Cp^{\circ}_{aq} = \Delta_h Cp^{\circ} + Cp^{\circ}_{ig}$, [h] Plyasunov *et al.* (2004), [i] calculated using $\Delta_h G^{\circ}$ and $\Delta_h H^{\circ}$ given in the table and the relation $\Delta_h G^{\circ} = \Delta_h H^{\circ} - T\Delta_h S^{\circ}$, [j] Domalski and Hearing (1993), [k] estimated from $\Xi_{C-(H)(C)_3} + \Xi_{C-(H)_3(C)} - 2 \times \Xi_{C-(H)_2(C)_2}$, where $\Xi_{C-(H)(C)_3}$, $\Xi_{C-(H)_3(C)}$, and $\Xi_{C-(H)_2(C)_2}$ are the properties of C-$(H)(C)_3$, C-$(H)_3(C)$, and C-$(H)_2(C)_2$ given in the table, respectively, [l] estimated from $\Xi_{C-(H)(O)(C)_2} + \Xi_{O-(H)(C)} - \Xi_{C-(H)_2(C)_2}$, where $\Xi_{C-(H)(O)(C)_2}$, $\Xi_{O-(H)(C)}$, and $\Xi_{C-(H)_2(C)_2}$ are the properties of C-$(H)(O)(C)_2$, O-(H)(C), and C-$(H)_2(C)_2$ given in the table, respectively, [m] estimated from $\Xi_{C-(H)_3(C)} - \Xi_{C-(H)_2(C)_2}$, where $\Xi_{C-(H)_3(C)}$ and $\Xi_{C-(H)_2(C)_2}$ are the properties of C-$(H)_3(C)$ and C-$(H)_2(C)_2$ given in the table, respectively.

Table 3.7: Partial molal thermodynamic data for pentacyclic and hexacyclic alkanes used in the estimation of GDGT internal ring properties.

	$\Delta_f G^{\circ}_{aq}$ [a]	$\Delta_f H^{\circ}_{aq}$ [a]	Cp°_{aq} [b]	V°_{aq} [c]	$\Delta_h G^{\circ}$ [a]
methylcyclopentane	51.16	-139.79	508	99.90	14.58
propylcyclopentane	69.93	-188.87	686	132.00	16.97
pentylcyclopentane	92.56	-233.71	855	163.40	19.06
1,1,3-trimethylcyclohexane	54.35	-252.15	794	158.00	17.14
$\Delta\Xi_{pent}$	34.27[d]	61.04[d]	-135[d]	-14.13[d]	-2.84[d]
$\Delta\Xi_{hex}$	8.11[e]	23.09[e]	-114[e]	-3.51[e]	-3.37[e]

Properties in this table are recommended values from the ORganic Compounds HYDration properties (ORCHYD) database compiled by Plyasunova *et al.* (2004) unless otherwise noted. [a] kJ mol^{-1}, [b] J mol^{-1} K^{-1}, [c] cm^3 mol^{-1}, [d] estimated by taking the average of $\Xi_{methylcyclopentane} - \Xi_{hexane}$, $\Xi_{propylcyclopentane} - \Xi_{octane}$, and $\Xi_{pentylcyclopentane} - \Xi_{decane}$, where $\Xi_{methylcyclopentane}$, $\Xi_{propylcyclopentane}$, and $\Xi_{pentylcyclopentane}$ are the properties of methylcyclopentane, propylcyclopentane, and pentylcyclopentane given in the table and Ξ_{hexane}, Ξ_{octane}, and Ξ_{decane} are the properties of hexane, octane, and decane given in Table 3.2, [e] estimated from $\Xi_{1,1,3-trimethylcyclohexane} - \Xi_{nonane}$, where $\Xi_{1,1,3-trimethylcyclohexane}$ is the properties of 1,1,3-trimethylcyclohexane given in the table and Ξ_{nonane} are the properties of nonane given in Table 3.2.

observed directly or quantified. Therefore, the value from nMethyl accounts only for methyl branching assumed to be present in isoprenoidally derived alkyl chains, and, as such, should be interpreted as a minimum because non-isoprenoidal chains may also have methyl branching. The contribution of $\Delta\Xi_{methyl}$ tends to be minor relative to those from other alkyl chain modifications, so I do not anticipate variation due to uncertainty in nMethyl to have a considerable effect on Ξ_{chain}. The term nMethyl was estimated from

$$\text{nMethyl} = (\text{nC}/5 - \text{nPent} - \text{nHex})(x_{GDGT} + x_{AR}), \tag{3.7}$$

where x_{AR} stands for the mole fraction of alkyl chains belonging to archaeol (AR), found in Table 3.1. The term nC/5 accounts for one methyl branch per five carbons in isoprenoidally derived alkyl chains, represented here as the sum of the mole fractions x_{GDGT} and x_{AR}. Because each pentacyclic and hexacyclic ring incorporated into the alkyl chains of GDGTs replaces a methyl group, nPent and nHex are subtracted when determining the number of methylations.

3.2.5 Calculation of alkyl$_{\text{free}}$ relative abundance at hot spring conditions.

The revised Helgeson-Kirkham-Flowers (HKF) equation of state (Shock *et al.*, 1992) was used to predict equilibrium constants for reactions at the temperatures measured at the hot spring sample locations. The HKF equation of state is shown in Appendix D, Equation D.1. Equation of state parameters a_1, a_2, a_3, a_4, c_1, c_2, and ω, for alkyl$_{\text{free}}$ are reported in Table 3.8 and were estimated according to the correlation strategies described in Plyasunov and Shock (2001) for aqueous polar organic nonelectrolytes.

Predictions of the relative abundances of aqueous free alkyl chains assumed thermodynamic equilibrium between the chain and a set of basis species consisting of

Table 3.8: Estimated thermodynamic standard state partial molal properties and Helgeson-Kirkham-Flowers equation of state parameters for aqueous sample-averaged free alkyl chains.

site	sample	formula	$\Delta_f G^\circ_{aq}$ [a]	$\Delta_f H^\circ_{aq}$ [a]	S°_{aq} [b]	Cp°_{aq} [b]	V°_{aq} [c]	a_1 [d] $\times 10$	a_2 [e] $\times 10^{-2}$	a_3 [f]	a_4 [g] $\times 10^{-4}$	c_1 [b]	c_2 [f] $\times 10^{-4}$	ω [e] $\times 10^{-5}$	$\Delta_h G^\circ$ [a]
Bison Pool	BP1	$C_{19.8}H_{39.5}N_{1.11e\text{-}2}O_{1.26}$	-20.3	-690.6	576.6	1670	317.3	260	200	380	-140	1600	17	-0.79	-4.8
	BP2	$C_{19.4}H_{38.6}N_{1.64e\text{-}2}O_{1.40}$	-55.8	-706.6	596.1	1640	311.1	260	190	370	-140	1600	17	-0.77	-5.3
	BP3	$C_{18.0}H_{35.6}N_{2.37e\text{-}2}O_{1.70}$	-146.4	-743.4	603.4	1510	290.6	240	180	340	-130	1500	16	-0.72	-6.6
	BP4	$C_{17.3}H_{34.8}N_{1.51e\text{-}2}O_{1.53}$	-112.7	-693.7	582.8	1460	280.2	230	170	330	-120	1400	15	-0.70	-7.4
	BP5	$C_{16.9}H_{33.1}N_{2.83e\text{-}3}O_{1.79}$	-167.4	-719.5	592.1	1410	272	230	170	320	-120	1400	15	-0.7	-7.4
	BP6	$C_{16.7}H_{31.8}N_{3.36e\text{-}3}O_{1.95}$	-180.5	-709.2	600.7	1380	265.1	222	159	302	-113	1350	12.8	-0.607	-10.1
Mound Spring	MS1	$C_{20.0}H_{39.8}N_{4.78e\text{-}5}O$	36.5	-654.9	498.9	1670	318.9	260	200	390	-140	1600	18	-0.81	-4.2
	MS2	$C_{19.6}H_{39.3}N_{4.74e\text{-}2}O_{1.12}$	0.3	-669.0	555.0	1660	314.6	260	200	380	-140	1600	18	-0.81	-4.0
	MS3	$C_{18.0}H_{36.3}N_{1.55e\text{-}2}O_{1.30}$	-46.8	-655.8	567.3	1520	289.7	240	180	340	-130	1500	16	-0.72	-6.7
	MS4	$C_{17.4}H_{34.3}N_{5.10e\text{-}3}O_{1.77}$	-160.5	-735.0	596.4	1460	280.7	230	170	330	-120	1400	15	-0.69	-7.5
	MS5	$C_{16.8}H_{31.7}N_{3.52e\text{-}3}O_{1.93}$	-161.7	-687.5	602.4	1380	264.7	222	158	300	-112	1350	12.5	-0.595	-10.5
Empress Pool	EP1	$C_{19.8}H_{39.8}N_{1.11e\text{-}3}O_{1.01}$	26.5	-661.2	511.5	1670	317.9	260	200	390	-140	1600	18	-0.81	-4.1
	EP2	$C_{18.9}H_{37.9}N_{3.15e\text{-}2}O_{1.20}$	-21.5	-664.6	553.8	1590	303.2	250	190	360	-130	1500	17	-0.76	-5.5
	EP3	$C_{19.1}H_{37.9}N_{1.07e\text{-}2}O_{1.27}$	-36.7	-687.6	534.3	1600	306.5	250	190	370	-140	1600	17	-0.78	-5.0
	EP4	$C_{17.8}H_{35.0}N_{5.91e\text{-}3}O_{1.63}$	-124.0	-712.4	583.5	1490	285.6	240	180	340	-120	1400	16	-0.73	-6.6
	EP5	$C_{17.1}H_{33.0}N_{7.42e\text{-}4}O_{1.76}$	-140.2	-690.2	590.5	1420	272.0	230	170	320	-120	1400	15	-0.68	-7.9
Octopus Spring	OS1	$C_{20.4}H_{41.3}N_{3.24e\text{-}3}O_{1.26}$	-23.3	-722.8	599.1	1750	329.7	270	210	410	-150	1700	18	-0.83	-3.5
	OS2	$C_{17.5}H_{35.3}N_{8.71e\text{-}3}O_{1.55}$	-116.0	-704.6	592.6	1480	283.7	240	170	330	-120	1400	15	-0.71	-7.1

[a] kJ mol^{-1}, [b] J mol^{-1} K^{-1}, [c] cm^3 mol^{-1}, [d] J mol^{-1} bar^{-1}, [e] J mol^{-1}, [f] J K mol^{-1} bar^{-1}, [g] J K mol^{-1}

bicarbonate (HCO_3^-), ammonium (NH_4^+), protons (H^+), electrons (e^-), and water (H_2O). The only exception was for calculations performed at the conditions of sample MS5 where ammonium was below the detection limit. Therefore, ammonium was replaced by nitrate (NO_3^-) as the nitrogen-bearing basis species.

Reactions from basis species were written for each aqueous free alkyl chain with a number of HCO_3^-, NH_3^+, H^+, and e^- as reactants and water as a product, such that

$$a \cdot HCO_3^- + b \cdot NH_4^+ + c \cdot H^+ + d \cdot e^- = alkyl + f \cdot H_2O, \tag{3.8}$$

where $alkyl$ represents one mole of the $\text{alkyl}_{\text{free}}$ (given in Table 3.8), and a, b, c, d, and f are stoichiometric reaction coefficients of the basis species required to balance the reaction. For example, the reaction to form 1 mole of the sample-averaged alkyl chain of sample BP1, $alkyl_{BP1}$, was written as:

$$\begin{aligned} 19.8 \cdot HCO_3^- &+ 0.0111 \cdot NH_4^+ \\ +135.9467 \cdot H^+ &+ 116.1467 \cdot e^- \\ &= alkyl_{BP1} + 58.14 \cdot H_2O. \end{aligned} \tag{3.9}$$

Sample-averaged free alkyl chains belonging to the same hot spring outflow channel (Bison Pool, Mound Spring, Empress Pool, and Octopus Spring) were grouped together for speciation calculations. Assuming metastable equilibrium, the activity of the i^{th} alkyl chain of an outflow channel, $[alkyl_i]$, was estimated using the relation

$$[alkyl_i] = e^{A^*/RT}, \tag{3.10}$$

where R stands for the gas constant, T indicates temperature, and A^* designates a modified reaction affinity. A^* is defined as:

$$A^* \equiv RT \ln K/Q^* \tag{3.11}$$

where K represents the HKF-calculated equilibrium constant for the reaction and Q^* is the modified reaction quotient composed only of the activities of the basis species. For the formation of sample-averaged alkyl chains from basis species,

$$Q^* = \frac{[H_2O]^f}{[HCO_3^-]^a[NH_4^+]^b[H^+]^c[e^-]^d} \tag{3.12}$$

for all samples except for MS5, where NH_4^+ is replaced by NO_3^-. Activities of NH_4^+ and NO_3^- were assumed to be equivalent to the molal concentrations of ammonia and nitrate reported in Chapter 2. Activities of protons were calculated directly from field pH measurements and HCO_3^- activity was calculated from DIC in equilibrium with aqueous carbonic acid species at the pH and temperature of each sample.

Metastable equilibrium activities, $alkyl_i\%$, were calculated for the i^{th} sample-averaged alkyl chain relative to others (expressed as a percent) within the same hot spring using the equation

$$alkyl_i\% = \frac{[alkyl_i]}{\sum [alkyl_i]} \cdot 100\%. \tag{3.13}$$

Thermodynamic analyses were performed in R with the package CHNOSZ (Dick, 2008). Calculations involving Equation 3.13 were performed using the mosaic function, allowing temperature- and pH-dependent speciation of basis species such that bicarbonate could fully or partially speciate into carbon dioxide or carbonate, and ammonium into ammonia, when calculating K and Q^*. In the case of sample MS5 where nitrate was used in place of ammonia, nitrate was not allowed to speciate.

The activity of electrons in Q^* was used as the independent variable when solving for $alkyl_i\%$, allowing metastable equilibrium activities of $\text{alkyl}_{\text{free}}$ to change as a

function of Eh. Note this study equates 'activities' and 'abundances' of alkyl chains when describing predicted distributions of lipids.

A resolution of 800 calculations across a range of Eh from -0.56 to -0.30 volts for samples in Bison Pool and Octopus Spring, -0.70 to -0.40 volts for samples in Mound Spring, and -0.50 to -0.26 volts for samples in Empress Pool was used. These Eh ranges were selected to contain changes in alkyl chain speciation across the temperatures and basis species concentrations of all samples within the outflow channel of a given hot spring.

Equations 3.8 through 3.13 were combined to calculate values of Eh from the relative abundances of alkyl chains at each sample location, using the basis species concentrations for the same location. Alkyl chain-predicted Eh is defined as the Eh value, in volts, that results in a local maximum in percent abundance at the temperature and basis species concentrations of its own sample (*e.g.*, the Eh that maximizes $alkyl_{BP2}\%$ at the temperature and HCO_3^-, NH_4^+, and H^+ concentrations measured at BP2). For samples without a local maximum in alkyl chain-predicted Eh, namely the first and last samples in a hot spring (*e.g.*, BP1 and BP6 for Bison Pool), a percent abundance cutoff of 95% was used.

A water sample was not collected for ES2 during the 2012 field season, so concentrations of bicarbonate and ammonium are unknown. Consequently, lipid-predicted Eh cannot be calculated for this sample.

3.2.6 Sensitivity analysis

Thermodynamic predictions of $\text{alkyl}_{\text{free}}$ percent abundance are sensitive to potential sources of analytical uncertainty. Sensitivity tests were generated using a Monte Carlo-style partial randomization of underlying lipid data as described in Chapter 2, where all LC-MS IPL peak areas were varied randomly to at most $\pm 30\%$, and

slopes of calibration curves were varied to a maximum of 0.01× and 100× yielding re-quantified IPL abundances over the course of 999 iterations. After each iteration, $alkyl_i\%$ values were re-calculated as before using a resolution of 400 instead of 800.

3.3 Results

3.3.1 Water chemistry and Eh

Micromolal concentrations of DIC (calculated as bicarbonate ion) were measured in all four springs and are shown in Table 3.9. The concentration of DIC was relatively invariate among samples of the same spring, except for Empress Pool, which showed a downstream decrease from 3.35 to 1.95 molal. Table 3.9 also gives Eh values calculated from dissolved inorganic species reported in Chapter 2, including Eh_{O_2/H_2O}, NO_2^-/NO_3^-, $Eh_{NH_4^+/NO_3^-}$, and $Eh_{HS^-/SO_4^{2-}}$ from the redox couples indicated by each subscript. Despite a downstream increase in dissolved oxygen and an increase in the ratio of oxidized to reduced dissolved inorganic species in all four hot springs, calculated Eh values increased strongly at Mound Spring and weakly at Bison Pool, and remained about the same between Octopus Spring samples. Empress Pool exhibited a downstream decrease in Eh, largely attributable to its rising pH from 5.78 at EP1 to 8.42 at EP5, which spans the greatest range in pH among the four hot springs. Across samples, Eh calculated from one redox couple tended to co-vary strongly with that of any other redox couple.

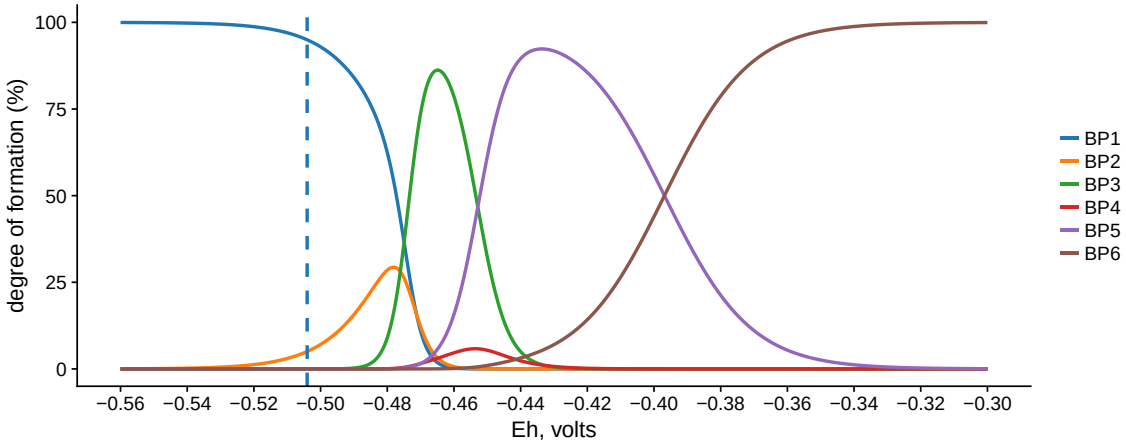
3.4 Relative equilibrium abundances of $alkyl_{free}$ and predicted Eh_{alkyl}

Figure 3.3 shows the relative equilibrium abundances predicted for $alkyl_{free}$ of Bison Pool as a function of Eh. Among all calculations performed at the temperatures and basis species concentrations of Bison Pool samples, the order in which

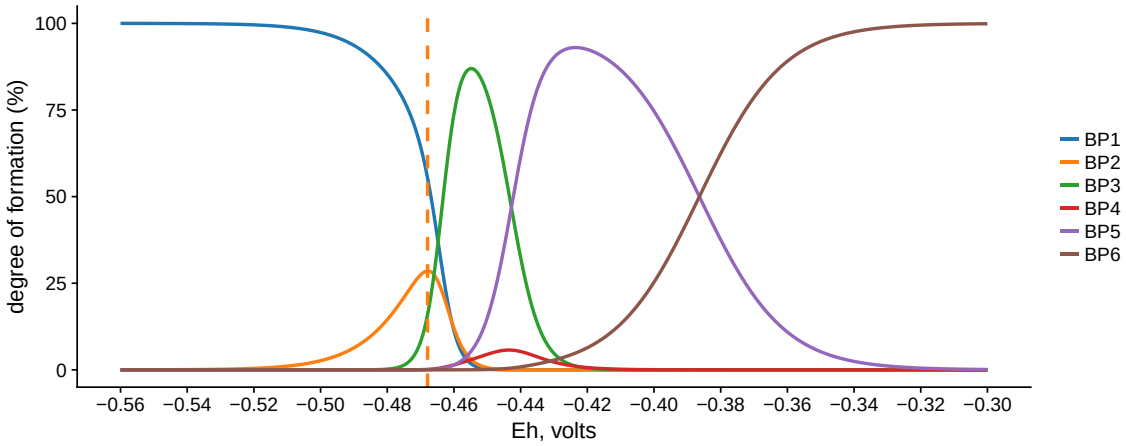
Table 3.9: Measured or calculated geochemical and physical data, along with Eh predicted from sample-averaged free alkyl chains.

Site	Sample	Temp. (°C)	pH	$[HCO_3^-]$ (mmolal)	Eh_{O_2/H_2O}[a] (volts)	$Eh_{NO_2^-/NO_3^-}$[b] (volts)	$Eh_{NH_4^+/NO_3^-}$[c] (volts)	$Eh_{HS^-/SO_4^{2-}}$[d] (volts)	Eh_{alkyl}[e] (volts)
Bison Pool	BP1	89.0	7.23	5.4	0.662	0.303	0.237	-0.294	-0.504
	BP2	80.9	7.34	5.4	0.679	0.295	0.242	-0.292	-0.468
	BP3	73.3	7.27	5.4	0.698	0.339	0.262	-	-0.435
	BP4	63.1	8.09	5.4	0.666	0.293	0.215	-0.316	-0.464
	BP5	40.5	8.25	5.22	0.707	-	0.242	-0.298	-0.394
	BP6	29.0	9.01	5.2	0.676	-	0.212	-0.335	-0.350
Mound Spring	MS1	91.0	8.81	2.8	0.550	-	0.101	-0.429	-0.658
	MS2	77.3	8.65	2.83	0.601	-	0.145	-0.397	-0.562
	MS3	64.8	9.08	2.9	0.592	0.228	0.094	-0.407	-0.557
	MS4	53.0	9.22	2.81	0.613	0.237	0.144	-0.395	-0.519
	MS5	35.1	9.53	3.0	0.634	-	-	-	-0.438
Empress Pool	EP1	82.2	5.78	3.35	0.781	0.392	0.367	-0.159	-0.383
	EP2	70.5	6.96	-	-	-	-	-	-
	EP3	60.7	7.63	2.24	0.696	-	0.242	-0.280	-0.451
	EP4	51.6	7.99	2.18	0.688	0.309	0.231	-0.295	-0.422
	EP5	38.1	8.42	1.95	0.692	0.328	0.229	-0.306	-0.393
Octopus Spring	OS1	85.4	7.29	5.38	0.672	0.314	0.243	-0.283	-0.467
	OS2	59.8	8.27	5.3	0.662	0.285	0.206	-0.328	-0.477

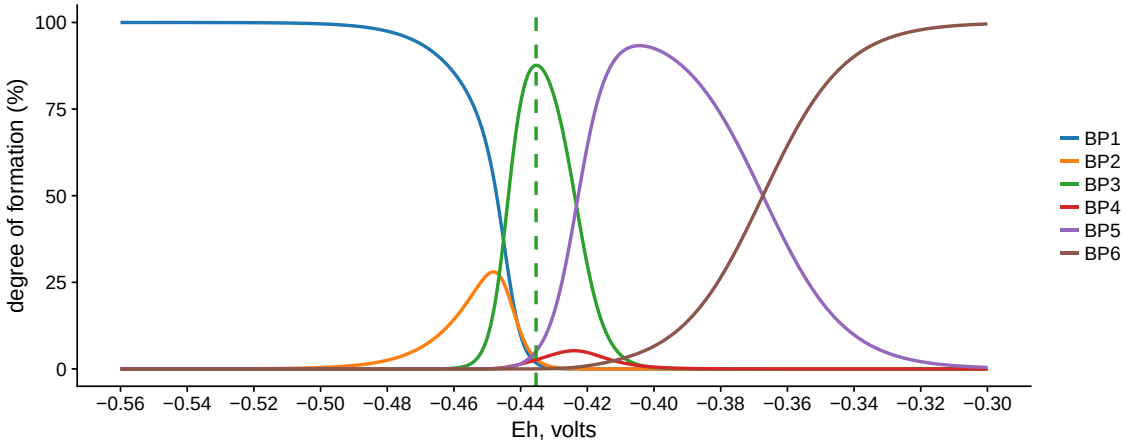
[a] Eh calculated from the redox couple $\frac{1}{2}O_2 + 2H^+ + 2e^- = H_2O$, [b] Eh calculated from the redox couple $NO_2^- + H_2O = NO_3^- + 2H^+ + 2e^-$, [c] Eh calculated from the redox couple $NH_4^+ + 3H_2O = NO_3^- + 10H^+ + 8e^-$, [d] Eh calculated from the redox couple $HS^- + 4H_2O = SO_4^{2-} + 9H^+ + 8e^-$, [e] Eh predicted from $alkyl_{free}$s as described in the text.


100
75
50
25
0
degree of formation (%)
−0.56 −0.54 −0.52 −0.50 −0.48 −0.46 −0.44 −0.42 −0.40 −0.38 −0.36 −0.34 −0.32 −0.30
Eh, volts
BP1
BP2
BP3
BP4
BP5
BP6

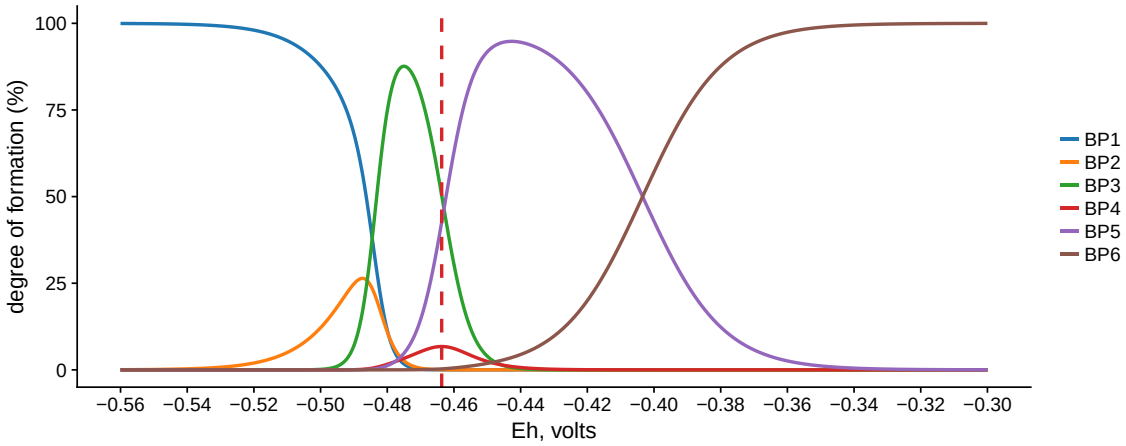
(a) BP1


100
75
50
25
0
degree of formation (%)
−0.56 −0.54 −0.52 −0.50 −0.48 −0.46 −0.44 −0.42 −0.40 −0.38 −0.36 −0.34 −0.32 −0.30
Eh, volts
BP1
BP2
BP3
BP4
BP5
BP6

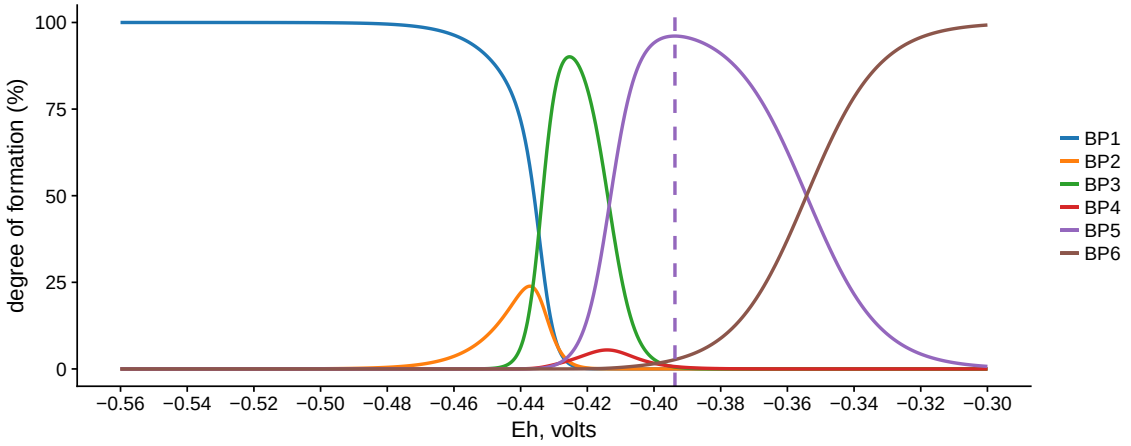
(b) BP2


100
75
50
25
0
degree of formation (%)
−0.56 −0.54 −0.52 −0.50 −0.48 −0.46 −0.44 −0.42 −0.40 −0.38 −0.36 −0.34 −0.32 −0.30
Eh, volts
BP1
BP2
BP3
BP4
BP5
BP6

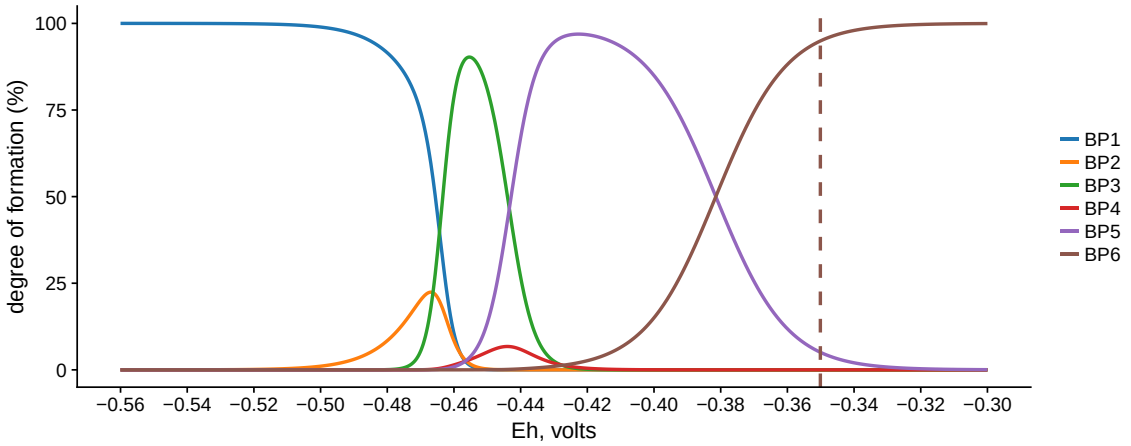
(c) BP3


100
75
50
25
0
degree of formation (%)
−0.56 −0.54 −0.52 −0.50 −0.48 −0.46 −0.44 −0.42 −0.40 −0.38 −0.36 −0.34 −0.32 −0.30
Eh, volts
BP1
BP2
BP3
BP4
BP5
BP6

(d) BP4



(e) BP5



(f) BP6

Figure 3.3: Metastable equilibrium abundance of sample-averaged free alkyl chains ($\text{alkyl}_{\text{free}}$) of Bison Pool predicted across an Eh gradient at the temperatures and bicarbonate, ammonium, and proton concentrations measured at (a) BP1, (b) BP2, (c) BP3, (d) BP4, (e) BP5, and (f) BP6. Vertical dotted lines indicate lipid-predicted Eh for the sample.

$alkyl_{free}$ were predicted to be stable relative to each other in Eh space did not change. In other words, the $alkyl_{free,\ BP1}$ was always most stable for the most reduced Eh values, $alkyl_{free,\ BP6}$ was always most stable under the most oxidized conditions, and $alkyl_{free,\ BP2}$ through $alkyl_{free,\ BP4}$ were always most stable in the same order. For each new set of basis species concentrations and temperature, the predicted abundance distributions were observed to shift together, rather than change shape in any significant way. The tendency to preserve the arrangement of predicted $alkyl_{free}$ abundance topography was also observed in calculations for Mound Spring (Figure B.1), Empress Pool (Figure B.2), and Octopus Spring (Figure B.3). The only exception was sample MS5 (Figure B.1e), which had significantly altered abundance topography due to its alternate set of basis species. This is discussed in more detail later.

In Figure 3.3a, at the temperature and basis species concentrations measured at BP1, the $alkyl_{free}$ representing observed alkyl chain distributions in BP1 predicted an Eh_{alkyl} of -0.504 volts at the 95% maximum abundance cutoff used for samples at the beginning and end of a hot spring sample series. In the conditions of BP2 (3.3b), a lower temperature and a different chemical composition shifted all $alkyl_{free}$ abundance distributions by approximately +0.01 volts from where they were in BP1. The predicted Eh_{alkyl} of BP2 was -0.468 volts based on the maximum predicted abundance of the site-averaged chain in the conditions of BP2.

Interestingly, several sample-averaged alkyl chains were never predicted to comprise a dominant proportion at any Eh. The chain representing BP2 is one such example; BP1 remains the most stable at the Eh_{alkyl} of BP2, with a predicted abundance ratio of approximately 55:25:20 BP1:BP2:BP3. This phenomenon appears to be predictable to some extent through examination of the abundance-weighted structural characteristics and Z_C of alkyl chains reported in Chapter 2. For instance, consider that the predicted abundance of $alkyl_{free,\ BP4}$ never reaches more than 10%

abundance in Bison Pool. According to Table 2.5, the alkyl chains of BP4 follow the typical downstream trends relative to adjacent samples BP3 and BP5, including a decrease in the number of aliphatic carbons and an increase in degree of unsaturation. However, the abundance-weighted alkyl chain Z_C in BP4 indicates that carbon in the alkyl chains of this sample is more oxidized than either BP3 or BP5, owing to a greater mole fraction of ether-linked alkyl chains observed in BP4 ($x_{ether} = 0.414$) than would be expected from downstream trends in alkyl chain-backbone linkage as evidenced by the adjacent upstream sample BP3 having an even smaller fraction ($x_{ether} = 0.217$). This greater abundance of ether-linked alkyl chains in BP4 results in diminished stability of $alkyl_{free,\ BP4}$ as predicted by thermodynamic equilibrium with basis species across Eh. Despite the diminished abundance of $alkyl_{free,\ BP4}$, predicted $Eh_{alkyl,\ BP4}$ always occurs between that of $Eh_{alkyl,\ BP3}$ and $Eh_{alkyl,\ BP5}$, as shown in Figure 3.3d. This suggests the overall combination of chain modifications observed at BP4 (*i.e.*, number of aliphatic carbons, unsaturations, linkage type, *etc.*) is responsible for the arrangement of $Eh_{alkyl,\ BP4}$ relative to others.

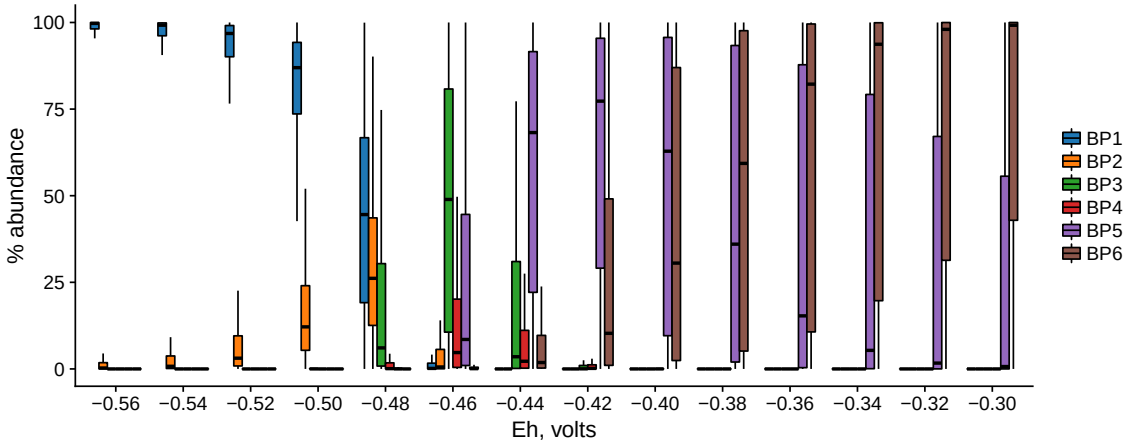
Concentrations of NO_3^- were used in place of NH_4^+ in the thermodynamic speciation of $alkyl_{free}$ at MS5. This alteration greatly extended the Eh range in which $alkyl_{free,\ MS2}$ was predicted to be maximally stable, and greatly diminished that of $alkyl_{free,\ MS1}$, compared to calculations using the basis species set of other Mound Spring samples (Figures B.1a through d in Appendix B). This can be explained by the relatively high concentration of dissolved NO_3^- measured in sample MS5 that would drive up the stability of $alkyl_{free}$ containing a greater proportion of nitrogen. Nitrogen is only introduced into the structure of $alkyl_{free}$ by the fractional contribution of a fatty amide functional group that is weighted by the mole fraction of observed amide-linked IPL alkyl chains (x_{amide}). Among Mound Spring samples, the smallest observed mole fraction of amide-linked chains was in MS1 ($x_{amide} = 4.78\text{e-}5$),

while the largest was in MS2 (x_{amide} = 4.47e-2), as reported in Chapter 1, Table 2.5. It follows that the relatively high concentration of NO_3^- used in the speciation of $alkyl_{free}$ in MS5 extended the Eh stability range of $alkyl_{free,\ MS2}$ owing to its higher nitrogen content relative to $alkyl_{free,\ MS1}$. Furthermore, this was only observed when using NO_3^- as a basis species and not NH_4^+ was due to the difference in the oxidation state of these inorganic molecules; NO_3^- is far more oxidized, so its limited incorporation into $alkyl_{free,\ MS1}$ diminished its Eh stability range relative to $alkyl_{free,\ MS2}$. This serves to illustrate the intriguing complexities that may arise when interpreting results from calculations employing a different choice of basis species.

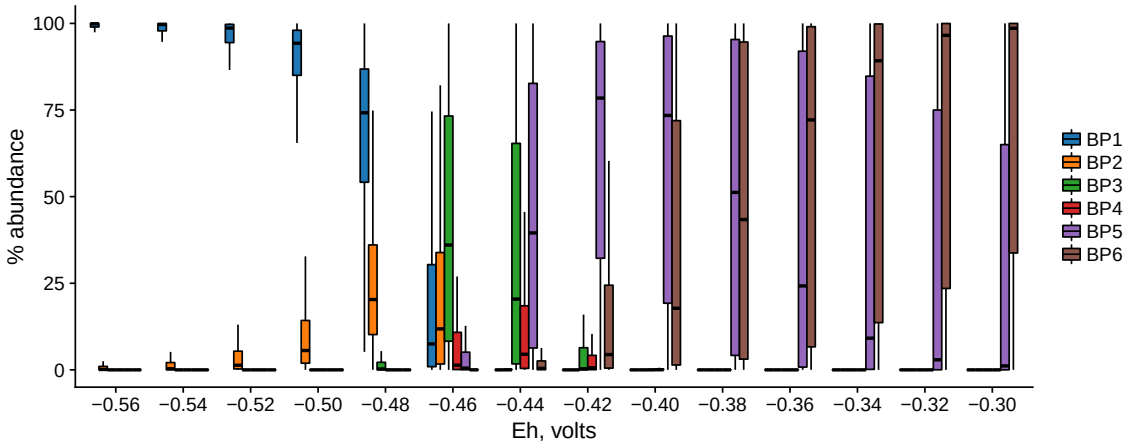
3.4.1 Sensitivity analysis of predicted equilibrium distributions

The results of the bootstrap sensitivity analysis indicate that relative $alkyl_{free}$ abundances predicted by thermodynamic calculations are altered by partial randomization in the analytical parameters used to quantify IPLs, but not to an extent that greatly changes the statistical distribution $alkyl_{free}$ topography across Eh. The 'canon' topographies of $alkyl_{free}$ for Bison Pool in Figure 3.3 remain recognizable within the distributions of relative abundances predicted over 999 iterations, as shown in figure 3.4. This was also the case for Mound Spring, Empress Pool, and Octopus Spring, as shown in Appendix B; compare Figures B.1 and B.4 for Mound Spring, Figures B.2 and B.5 for Empress Pool, and Figures B.3 and B.6 for Octopus Spring. This suggests the thermodynamic calculations performed in this study are resistant to sources of analytical error and persist despite the semiquantitative nature of the HPLC-MS method used to calculate abundance-weighted structural properties of alkyl chains.

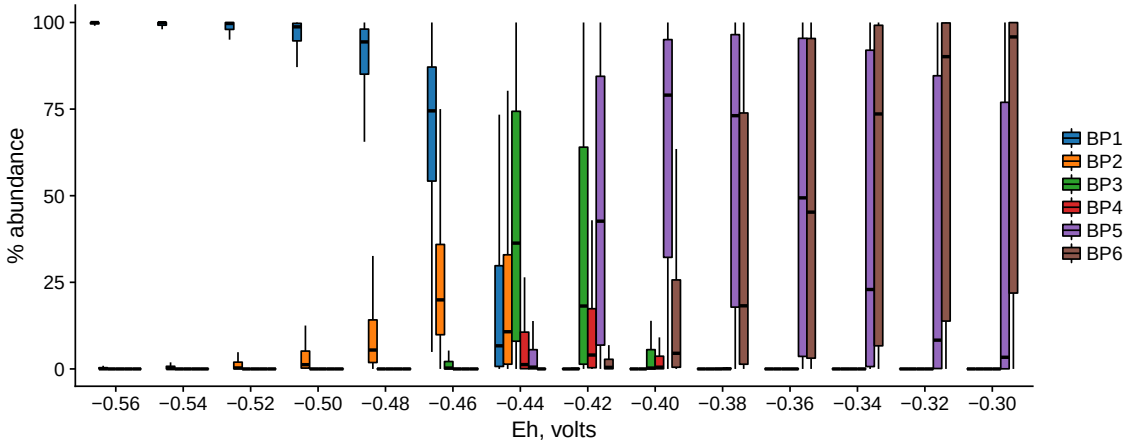
As shown in Figure 3.5, there is a weak positive correlation between Eh_{alkyl} and temperature across all samples (R^2 = 0.35). The correlations tended to be stronger among samples within the same outflow channel for Bison Pool and Mound Spring.


100
75
50
25
0
% abundance
−0.56
−0.54
−0.52
−0.50
−0.48
−0.46
−0.44
−0.42
−0.40
−0.38
−0.36
−0.34
−0.32
−0.30
Eh, volts
BP1
BP2
BP3
BP4
BP5
BP6

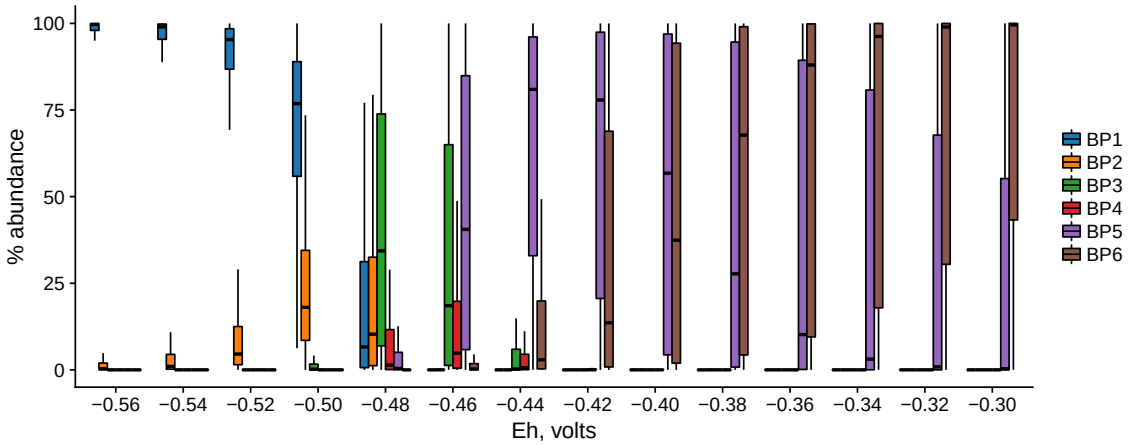
(a) BP1


100
75
50
25
0
% abundance
−0.56
−0.54
−0.52
−0.50
−0.48
−0.46
−0.44
−0.42
−0.40
−0.38
−0.36
−0.34
−0.32
−0.30
Eh, volts
BP1
BP2
BP3
BP4
BP5
BP6

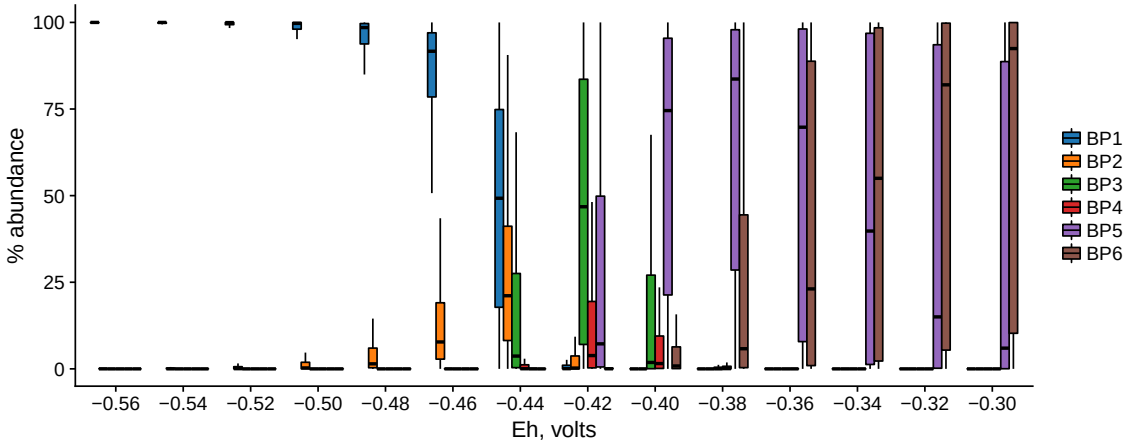
(b) BP2


100
75
50
25
0
% abundance
−0.56
−0.54
−0.52
−0.50
−0.48
−0.46
−0.44
−0.42
−0.40
−0.38
−0.36
−0.34
−0.32
−0.30
Eh, volts
BP1
BP2
BP3
BP4
BP5
BP6

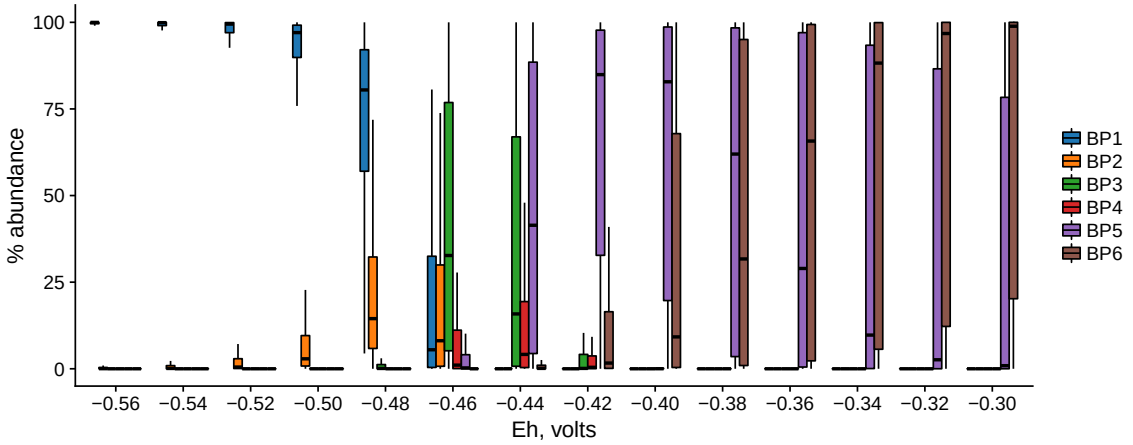
(c) BP3


100
75
50
25
0
% abundance
−0.56
−0.54
−0.52
−0.50
−0.48
−0.46
−0.44
−0.42
−0.40
−0.38
−0.36
−0.34
−0.32
−0.30
Eh, volts
BP1
BP2
BP3
BP4
BP5
BP6

(d) BP4



(e) BP5



(f) BP6

Figure 3.4: Sensitivity analysis of sample-averaged free alkyl chain metastable equilibrium abundance in Bison Pool samples (a) BP1, (b) BP2, (c) BP3, (d) BP4, (e) BP5, and (f) BP6 after 999 iterations of Monte Carlo-style random variation of lipid HPLC peak areas by 30% and mass spectral response factors between 0.01x and 100x. Results have a resolution of 400 calculations per iteration, binned here in 0.02 volt increments. Dark horizontal lines indicate median values for each distribution, with 50% of the results around the median falling within the colored box (interquartile range, or IQR). Whiskers extend to an observation 1.5 times the IQR beyond this range. Values that fall outside the span of the whiskers are not indicated to reduce visual clutter.

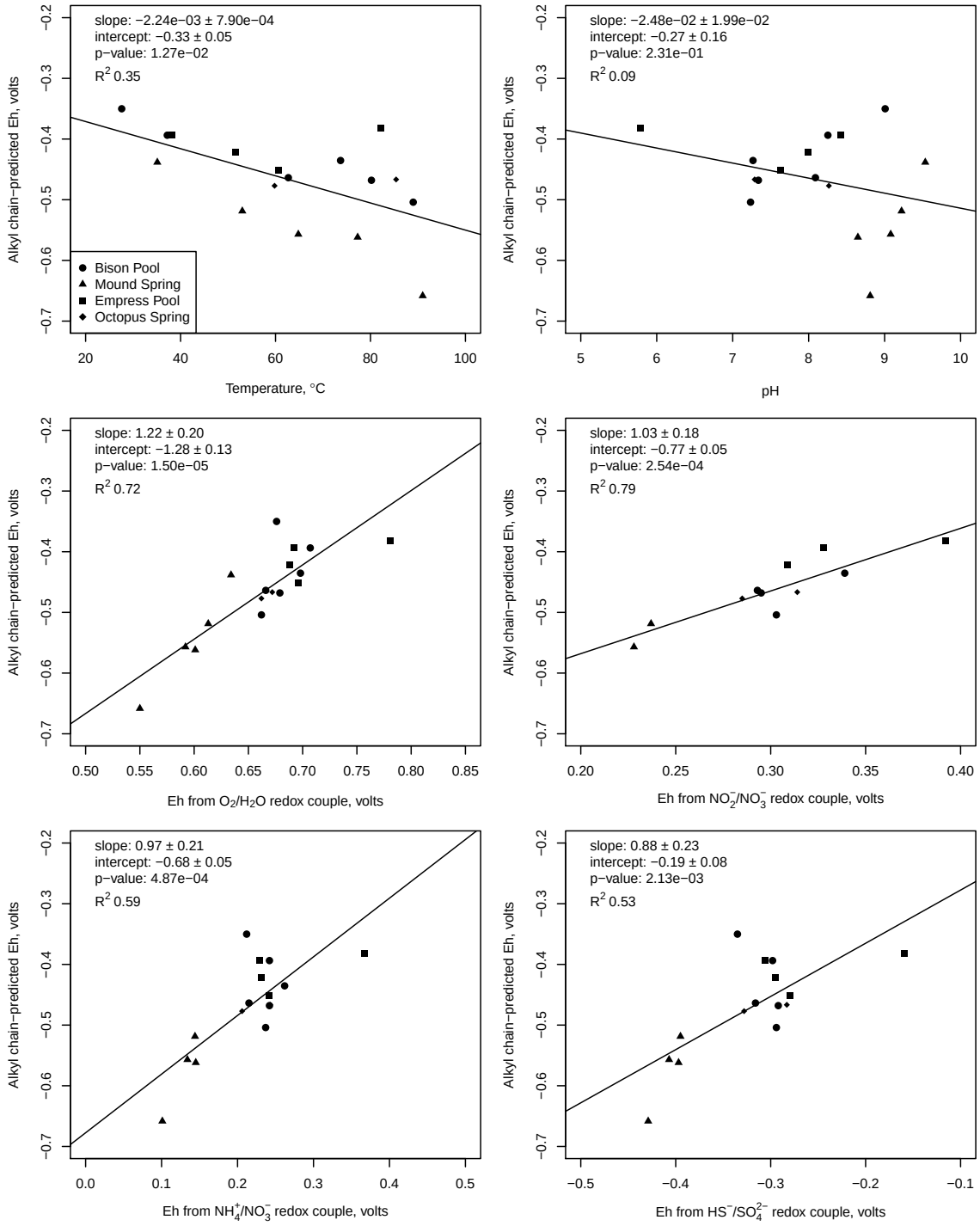


Figure 3.5: Alkyl chain-predicted Eh as a function of temperature, pH, and Eh.

A correlation was not observed between pH and Eh_{alkyl} ($R^2 = 0.09$).

Positive correlations were found between Eh_{alkyl} and Eh calculated from various redox couples reported in Table 3.9; $Eh_{NH_4^+/NO_3^-}$ ($R^2 = 0.63$), and $Eh_{HS^-/SO_4^{2-}}$ ($R^2 = 0.57$). The strongest correlations were found between Eh_{alkyl} and Eh calculated from $Eh_{NO_2^-/NO_3^-}$ ($R^2 = 0.87$) and Eh_{O_2/H_2O}, ($R^2 = 0.72$). $Eh_{NO_2^-/NO_3^-}$ was only calculated for eleven of the seventeen available water samples because the remaining samples had NO_3^- concentrations below the detection limit. Therefore, I posit that the correlation between Eh_{O_2/H_2O} that involves all available values in this sample set, may represent a more meaningful correlation with Eh_{alkyl} and is shown in Figure 3.5.

3.5 Discussion

Alkyl chains of IPLs sampled from microbial communities across the temperature and redox gradients of four alkaline hot springs in YNP with reduced structural modifications were most abundant under hot, reducing upstream conditions, while those with oxidized modifications were most abundant under oxidized conditions downstream (Chapter 2). It was hypothesized that these changes were the result of adaptation to the temperature and geochemistry of the system and that alkyl chain modifications trended with redox gradient in an energetically favorable way. The results of this study show that simplified molecules representing entire lipid distributions, $alkyl_{free}$, are predicted to speciate according to chemical equilibrium along redox gradients. The formation of $alkyl_{free}$ representatives of lipid distributions sampled from upstream hot, reducing conditions, were likewise favored under reduced Eh regimes. The $alkyl_{free}$ values for of downstream sites were predicted to be most stable under oxidized conditions. Furthermore, the Eh values that predicted maximal abundances of $alkyl_{free}$ were always in the order of most reduced to most oxidized for upstream to downstream samples, respectively. This suggests the upstream

to downstream change observed in IPL distributions not only function to provide thermostable membranes, but also fulfill this function by incorporating energetically cost-effective structural changes into alkyl chains along an Eh gradient. This appears to corroborate trends in abundance-weighted Z_C for IPL-linked alkyl chains reported in Chapter 2. However, this analysis also revealed subtleties that could not be predicted by Z_C alone. The IPL-linked alkyl chains of sample BP4 were out-of-order with respect to Z_C trends downstream of Bison Pool, but contained a combination of lipid structural modifications that placed its maximum predicted abundance at an Eh between those of BP3 and BP5 rather than an out-of-order Eh. This suggests Z_C is a useful but imperfect proxy for general trends observed in the equilibrium speciation calculations of $\text{alkyl}_{\text{free}}$.

Abundance-weighted Z_C implied that carbon in IPL headgroups is becoming more reduced, rather than more oxidized, downstream in all four hot springs (Chapter 2). This trend was sensitive to simulated sources of analytical uncertainty and was proposed to be an artifact of the HPLC-MS method used to quantify IPL headgroups. However, if this is not the case and carbon in IPL headgroups does, in fact, become more reduced downstream, then it is worth exploring possible ramifications for energetic cost Equilibrium abundance calculations involving lipid headgroups serve as the focus of Chapter 4.

Throughout this study, calculations assumed equilibrium between sample-averaged free alkyl chains. Equilibrium thermodynamic models cannot account for every energetic necessity in biology; for instance, a recent study by Amenabar *et al.* (2017) showed a strain of the archaeon *Acidianus* grew preferentially with an alternate set of growth substrates than those predicted to yield the most metabolic energy. However, thermodynamic predictions may have a better chance of describing general trends in energetically driven biomolecular or metabolic adaptations to environmental

conditions of entire microbial communities rather than in individual organisms.

This work employed the use of free, rather than linked, alkyl chains, but future work could attempt similar predictions of equilibrium speciation with hybrid ether/ester/C-C/amide-bonded sample-averaged alkyl chains ($\text{alkyl}_{\text{linked}}$) to determine whether there are any major differences in predicted stability with the free chains of this work. At high temperatures, such as those in the upstream samples of hot springs, it is fathomable that an ether-linked chain is more stable relative to a hydrolyzed free fatty alcohol; perhaps more so than an ester-linked chain relative to a hydrolyzed fatty acid at those same conditions. This could provide thermodynamic evidence that ethers are preferable to esters at high temperatures; *i.e.*, that less of a thermophile's energy budget would need to be diverted to re-attaching hydrolyzed alkyl chains. Future work could involve the estimation of ester-, ether-, C-C, and amide-linked sample-averaged chains to investigate whether predicted equilibrium abundances differ significantly from those of free chains. Calculations across temperature, pressure, and pH might offer insight into the conditions where chain linkage types are most susceptible or resistant to hydrolysis. Low-pH conditions could be explored in this way to investigate potential temperature-pH-redox interplay in the thermodynamic stabilities of sample-averaged ethers, esters, and hydrolyzed chains to gain insight into geochemical controls on acidophile membrane homeostasis.

3.6 Conclusion

Alkyl chains from observed IPLs were used to estimate the thermodynamic properties of aqueous site-averaged free alkyl chains. These hypothetical molecules represent the abundance-weighted structural characteristics of a sample such as number of aliphatic carbons, degree of unsaturation, isoprenoidal branching, *etc.*. Thermodynamic calculations predicted that alkyl chains were most stable and potentially

energetically cost-effective for microbial communities to produce when distributed along a redox gradient. Under more reduced conditions, the alkyl chains representing lipids extracted from the hot, reduced condition upstream samples were predicted to be most energetically cost-effective, and *vice versa* for low-temperature oxidized conditions and downstream lipids. These results provide evidence that the weighted Z_C is a proxy for energetic favorability in lipid alkyl chains. Further, the Eh values predicted to maximize the stabilities of site-averaged free alkyl chains under the temperature and chemical compositions of their own sample are correlated positively with Eh values calculated from inorganic redox couples. This supports the hypothesis that the external environment is driving lipid adaptation, not only to produce thermostable membranes but also to minimize energetic cost.

Chapter 4

THERMODYNAMIC ASSESSMENT OF REDOX CONTROLS ON BACTERIOHOPANEPOLYOL HEADGROUPS

4.1 Introduction

Trends in the average oxidation state of carbon (Z_C) in intact polar lipids (IPLs) extracted from microbial communities sampled along four hot spring outflow channels in Yellowstone National Park (YNP) indicate that these molecules become more oxidized downstream, corresponding to decreasing temperature and increasingly oxidized conditions, as discussed in Chapter 2. I also showed downstream changes in alkyl chain structure were most responsible for trends in Z_C of full IPLs. Alkyl chain adaptations that had the largest impact on Z_C were: shifts in ether- to ester-linkage, shortened chain length, increased degree of unsaturation, and increased degree of glycerol dialkyl glycerol tetraether (GDGT) cyclization. Each of these structural modifications contributed to an increase in lipid carbon oxidation downstream. Because of this, I argued adaptations generally thought to maintain membrane function along thermal gradients are also adaptations to the chemical gradients of these systems. Specifically, the oxidation-reduction (redox) potential was invoked as a major driver in evolutionary convergence on lipid modifications that provide biological function and are also energetically favorable. The reasoning was, if there are several possible alkyl chain modifications that solve the same problem (*e.g.*, providing thermotolerant membranes), the one selected by evolution is the one that does the best job for the smallest energetic cost. I posited that oxidized biomolecules are energetically stable and cost-efficient under oxidized conditions and would explain the abundance of oxi-

dized chain modifications in the oxidized, low-temperature samples. Similarly, it may explain why lipids adapted to the hot, reduced conditions near the hot spring source had the most reduced alkyl chain modifications.

Carbon in another major component of IPLs, the headgroup, did *not* become more oxidized downstream, as reported in Chapter 2. Instead, the opposite trend was observed, with the weighted Z_C of IPL headgroups seemingly indicating that carbon was becoming more reduced with decreasing temperature and increasingly oxidized conditions. While the results of the sensitivity analysis indicated that this trend might be an artifact of the semiquantitative methods used to analyze IPL abundance in complex environmental samples, it is worth entertaining the possibility that carbon in lipid headgroups is becoming more reduced downstream in this sample set. This raises the question: would relatively reduced carbon in lipid headgroups downstream, or relatively oxidized carbon in headgroups upstream, translate to an associated increase in energetic cost? It would seem so based on the stated hypothesis that oxidized and reduced biomolecules are energetically favorable to produce under oxidized and reduced conditions, respectively. Therefore, the study represented here was conducted to explore the possibility that a more rigorous thermodynamic treatment might reveal evidence that observed trends in lipid headgroups are in an energetically favorable trajectory with respect to temperature and chemical composition. In this way, I sought to address the hypothesis that **adaptations to maximize the function of lipid headgroups also minimize cost**.

Attempts to perform thermodynamic estimations for IPL headgroups presented a series of challenges. For one, a complex variety of IPL headgroups was reported for hot spring samples in Chapter 2. The chemical formulae and, in most cases, basic structural characteristics of these headgroups were determined analytically or referenced from the literature. However, certain structural details important for thermodynamic

predictions could not be obtained, such as major carbohydrate configurations (*e.g.*, whether a hexose moiety was an inositol or an aldohexose sugar such as glucose). The chemical structure of one headgroup, nicknamed '223', could not be determined at all. Another challenge facing thermodynamic estimation of IPL headgroups was the scarcity of experimentally derived thermodynamic data needed to estimate the partial molal properties of certain IPL headgroup structures. For instance, thermodynamic measurements of aqueous organic molecules bearing sulfonic acid groups or quaternary amine groups were rare to nonexistent in my literature search. Without properties of aqueous sulfonic acids, the properties of aqueous sulfoquinovosyl (SQ) headgroups common in low-temperature samples could not be estimated. Similarly, the properties of phosphatidylcholine (PC) and several other aminolipids and phospholipids could not be estimated without literature values for a quaternary amine. It was decided that these challenges could be circumvented with a dataset containing a smaller variety of lipid headgroups with thermodynamic properties more easily estimated from experimental data available in the literature.

The lipids chosen for this study were bacteriohopanepolyols (BHPs), represented in Figure 4.1. BHPs are 'C_{35} hopanoids' (*i.e.*, hopanoids containing 35 carbons) produced by a variety of bacteria (Ourisson and Rohmer, 1982; Rohmer *et al.*, 1984). Hopanoids have a specialized 5-ringed hydrophobic region and have been found to thicken and condense bacterial membranes, as demonstrated first by Poralla *et al.* (1980). A later experiment by Poralla *et al.* (1984) showed hopanoid abundance in *Bacillus acidocaldarius* was strongly correlated with temperature, suggesting hopanoids play a role in membrane thermotolerance. Welander *et al.* (2009) observed defects in mutant cultures of *Rhodopseudomonas palustris* strain TIE-1 lacking BHPs when grown under acidic or alkaline conditions, suggesting these lipids may impart pH resistance. A study with a mutant *Burkholderia cenocepacia* also showed

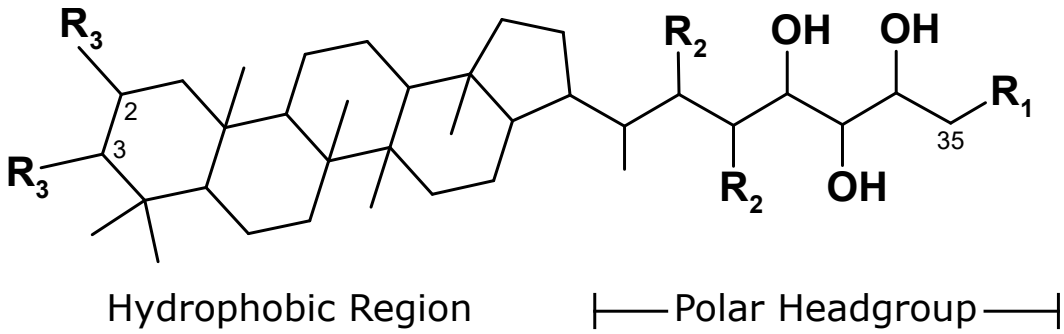


Figure 4.1: Generalized bacteriohopanepolyol (BHP) chemical structure. For BHPs without composite groups, R_1 represents a hydroxyl (OH), amine (NH_2), or amminium (NH_3^+) group. R_2 stands for H or OH. The BHP is considered pentafunctional if one R_2 is OH, or hexafunctional if both are OH. R_3 stands for H or CH_3. Carbon positions 2, 3, and 35 are indicated.

defects in the organism at low pH when unable to produce hopanoids, as well as poor swimming motility and decreased resistance to antibiotics and detergents (Schmerk *et al.*, 2011). Another study involving cultures of ethanol-producing *Zymomonas mobilis* concluded that hopanoid-fatty acid interaction conferred resistance to high ethanol concentrations (Bringer *et al.*, 1985). Examination of nitrogen-fixing vesicles in *Frankia* strains showed most lipids belonged to two BHP structures that are likely responsible for oxygen impermeability in these membranes (Berry *et al.*, 1993). A study with *Streptomyces coelicolor* found that hopanoid expression may help prevent the escape of cytoplasmic water during aerial growth by condensing membranes (Poralla *et al.*, 2000). Together, these studies suggest that BHPs fulfill a role analogous to sterols in Eukaryotic organisms by regulating the fluidity and permeability of lipid membranes.

Hopanoids are produced from the cyclization of squalene into diploptene by the enzyme squalene-hopene cyclase, SHC. Diploptene can be further modified by a series of enzymes to gain a polyfunctionalized aliphatic side-chain (Belin *et al.*, 2018; Welander *et al.*, 2012), referred to in this work as a polar headgroup. BHPs are thought to orient in the membrane with the hydrophobic region facing inward and the polar headgroup facing outward, a notion supported by molecular dynamics simulations of bacteriohopanetetrol in a lipid bilayer (Poger and Mark, 2013).

The hydrophobic region of BHPs tends to show little structural variation in living organisms. In most reports, a methylation may appear at either the C2 or C3 carbon position (see Figure 4.1), or not at all. These methylations tend to be preserved in hopanoid structures in the geologic record. For this reason, there has been much study and discussion about the usefulness of hopanes as biomarkers (see review by Newman *et al.*, 2016).

BHP headgroups exhibit a variety of structures. The C35 terminus of the head-

group can be "decorated" with a hydroxyl group or an amine. The number of headgroup hydroxylations can vary, resulting in 4, 5, or 6 total functional groups. Other organic molecules called 'composite groups', such as cyclitol and glucosamine amino sugars, can link to the headgroups of BHPs. Intriguingly, the specific functions served by BHP headgroup structures are still generally undiscovered (Belin *et al.*, 2018).

In this study, I quantified structures of BHPs in thermophilic microbial communities sampled along the thermal and redox gradients of four YNP hot springs: Bison Pool (BP), Mound Spring (MS), Empress Pool (EP), and Octopus Spring (OS). To simplify subsequent thermodynamic calculations, I separated quantities of BHP headgroups into groups by degree of functionality (*i.e.*, tetra-, penta-, and hexafunctional) and by aminopolyol and polyol. BHP composite groups were not considered in this study. I estimated the aqueous thermodynamic properties of the six resulting BHP headgroups: tetrol, pentol, hexol, aminotriol, aminotetrol, and aminopentol, as well as the properties of the ionized forms of the aminopolyols, including amminiumtriol, amminiumtetrol, and amminiumpentol. Estimation of BHP headgroup aqueous thermodynamic properties were performed independently of observations of BHPs in hot spring samples and were based purely on thermodynamic measurements of organic molecules reported in the literature. Estimated BHP headgroup properties were used in conjunction with measurements of hot spring temperature and chemical composition to predict equilibrium speciation of aqueous BHP headgroups along reduction potential (Eh) gradients, following the same principles one would use to predict aqueous carbonate speciation along pH gradients. Relative abundance of BHP headgroups was then compared to these independent thermodynamic predictions of aqueous BHP headgroup speciation. Much like using measured abundances of carbonate species to predict pH, measured abundances of tetrafunctional and pentafunctional BHP headgroups were used to predict Eh in each sample. These $Eh_{BHP,\ head}$ values were com-

pared to values of Eh previously reported for this sample set (Chapter 3) and were found to have a strong positive linear correlation, especially those calculated from the O_2/H_2O redox couple.

Owing to the dominance of pentafunctional and tetrafunctional headgroups among samples, I predicted $Eh_{BHP,\ head}$ based on the ratio of summed pentafunctional to tetrafunctional BHP headgroups. I noted that thermodynamic predictions appeared to significantly overestimate the abundance of trace hexafunctional BHPs. Conversely, observed abundances of aminopolyols were underestimated by thermodynamic predictions by 1-2 orders of magnitude. This opens the door to future work evaluating the sources of these discrepancies between predicted and observed BHP abundances and whether they stem from analytical uncertainty, underlying thermodynamic assumptions, or biological membrane functional requirements necessitating a higher energy cost.

4.2 Methods

4.2.1 Water chemistry

Methods used to measure, collect, and analyze water samples are described in Chapter 2 and 3. Briefly, temperature was measured with a YSI 30 conductivity meter and pH with a model 3300i or 3110 WTW pH meter with WTW probe. A Hach model 2400 or 2800 portable spectrophotometer and Hach reagents/protocols were used to measure concentrations of dissolved oxygen and sulfide in unfiltered water samples in the field. Samples for laboratory analysis were filtered to 0.2 μm polyethersulfone (PES) syringe filters with Supor^{TM} (Pall Corporation) filters and collected in 30 mL HDPE Nalgene bottles for ion chromatography (IC) and acid-washed amber glass vials with black butyl rubber septa for analysis of dissolved inorganic carbon (DIC).

Concentrations of major cations and anions were determined on Dionex DX-600 IC systems, and DIC concentrations were measured on an OI brand Wet Oxidation Total Organic Carbon analyzer coupled to a Thermo Delta Plus Advantage isotope ratio mass spectrometer.

4.2.2 Analysis of BHPs

Total lipid extracts (TLEs) from hot spring sediments and biofilms were obtained for this sample set as described in Chapter 2. Briefly, samples were collected with sterilized forceps and spatulas, placed in sterile specimen containers, and stored on dry ice the field and transferred to a -80°C freezer in the laboratory. Samples were freeze-dried and homogenized with mortar and pestle before extraction using a modified version of the Bligh and Dyer method (White and Ringelberg, 1998).

Aliquots of TLE set aside for BHP analysis were prepared for HPLC-MS as described in Talbot *et al.* (2003a). Briefly, BHPs were converted into their acetate derivatives by heating aliquots of TLE in 1:1 (v/v) acetic anhydride and pyridine at 50° for 1 hour, leaving at room temperature overnight, drying under N_2, and then redissolving in methanol. Resulting TLEs were injected on an Agilent 1200 Series HPLC equipped with a binary pump linked to an Agilent Q-TOF 6520 MS with a Poroshell 120 EC-C18 column and atmospheric pressure chemical ionization (APCI) source operated in positive ion mode. The column was first eluted isocratically with 100% solvent A (95:5 MeOH:water) for 2 min, followed by a linear gradient to 20% solvent B (isopropanol) for 18 minutes, then held isocratically for 10 min. The linear gradient continued to 30% B over 10 min, held for 5 min, increased to 80% B over 1 min, then maintained for 14 min. The column was then eluted with 100% A for 5 minutes. The flow rate throughout analysis was 0.19 mL min^{-1}.

BHPs were identified based on their relative retention times, the exact masses of

their protonated molecular ion peaks at the MS level, and by referencing MS/MS fragmentation patterns to published spectra (Talbot *et al.*, 2005, 2007a,b, 2003a,b, 2008). Diagnostic fragments m/z 191.18 and 205.20 indicating hopanoid C-ring cleavage were used to identify nonmethylated and methylated hopanoids, respectively. Tetra-, penta-, or hexafunctional BHPs were identified by fragments corresponding to the monoisotopic mass of acetylated polyol BHPs after loss of an acetyl or composite group fragment, namely: m/z 655.49, 713.50, or 771.50 for polyol BHPs without a 2- or 3-methylation, and m/z 669.51, 727.51, or 785.52 for polyol BHPs with a 2- or 3-methylation. Aminopolyols were identified based on the m/z of the parent ion plus a proton adduct ($[M + H]^+$), and fragments at the MS/MS level indicating acetyl group loss. Aminopolyols with composite groups were not observed. Varying abundances of adenosylhopane and ribosylhopane were observed across samples but were not included in this study as their headgroup functionality was not readily comparable to other BHPs, and because they might serve mainly as intermediates in BHP biosynthesis (Bradley *et al.*, 2010; Liu *et al.*, 2014). Anhydrobacteriohopanepolyols were observed but were not included in this analysis as they are thought to be mainly BHP degradation products (Talbot *et al.*, 2005; Schaeffer *et al.*, 2008). Relative BHP within-sample abundances were obtained semi-quantitatively by comparing ratios of integrated peak areas of parent ions.

4.2.3 Estimation of standard state thermodynamic properties of BHP polar headgroups.

Thermodynamic properties for BHP headgroups were estimated assuming the species were in the aqueous phase. These include the aqueous partial molal thermodynamic Gibbs free energies, $\Delta_f G^\circ_{aq}$, and enthalpies of formation, $\Delta_f H^\circ_{aq}$, from the elements, volumes V°_{aq}, and heat capacities (Cp°_{aq}). Additionally, $\Delta_h G^\circ_{aq}$ was esti-

mated for each BHP headgroup, representing the Gibbs free energy of the hydration process. Standard state conventions for aqueous species, liquid water, and hydration properties are described in Chapter 3.

The workflow used to estimate standard state partial molal thermodynamic properties of aqueous BHP headgroups is shown schematically in Figure 4.2. Estimations of tetra-, penta-, and hexafunctional BHP headgroups began with the properties of 1,2,3,4-butanetetrol, 1,2,3,4,5-pentanepentol, and 1,2,3,4,5,6-hexanehexol, respectively (structure A). The properties of these three sugar alcohols were estimated by gathering available properties reported for aqueous sugar alcohol diastereomers (compiled in Table 4.1 along with their literature sources) and then regressing available $\Delta_f G^\circ_{aq}$, $\Delta_f H^\circ_{aq}$, Cp°_{aq}, V°_{aq}, and $\Delta_h G^\circ_{aq}$ as a function of carbon length. These regressions, shown in Figure 4.3, were used to solve for the properties of four, five, and six carbon-long isomer-agnostic sugar alcohols (given in Table 4.2) used as structure A in the BHP headgroup estimation workflow.

Steps one, two, and three of the workflow shown in Figure 4.2 involve the addition and subtraction of second order group contributions to thermodynamic properties given in Table 4.3 to modify the properties of sugar alcohols (structure A) into those estimated for non-amino BHP headgroups (structure D). In step 1, the properties of structure B are obtained from structure A by subtracting the contribution from C-$(H)_2$(O)(C), the properties of a CH_2 group bonded to one carbon and one oxygen, and adding the contribution from C-(H)(O)$(C)_2$, the properties of a CH group bonded to two carbons and one oxygen. In step 2, the contribution of C-$(H)_2(C)_2$, a CH_2 group bonded to two carbons, is added to structure B a number of times such that structure C has six carbons. Step 2 is skipped for hexafunctional structure B, as it already has six carbons. In step 3, contributions from C-(H)$(C)_3$ and C-$(H)_3$(C) are then added to structure C (or structure B for the hexafunctional workflow), resulting

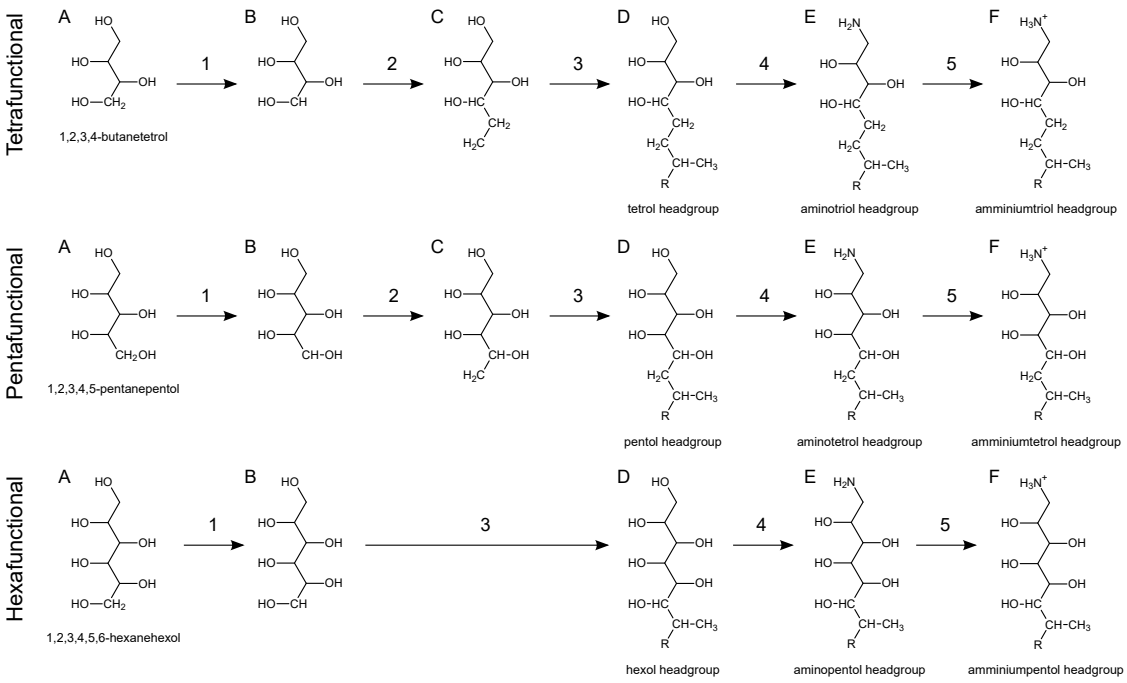


Figure 4.2: Workflow used in the estimation of partial molal standard state thermodynamic properties of aqueous BHP headgroups. Structures and steps are described in the methods. R stands for the rest of the BHP structure, though its thermodynamic properties are not considered by this study.

Table 4.1: Partial molal thermodynamic data for sugar alcohols used in the estimation of BHP polar headgroup properties

	nC	$\Delta_f G^{\circ}_{aq}$ [a]	$\Delta_f H^{\circ}_{aq}$ [a]	Cp°_{aq} [b]	V°_{aq} [c]	$\Delta_h G^{\circ}$ [a]	$\Delta_h H^{\circ}$ [a]	$\Delta_{sol} G^{\circ}$ [a]	$\Delta_{sol} H^{\circ}$ [a]	$\Delta_{sub} G^{\circ}$ [a]	$\Delta_{sub} H^{\circ}$ [a]	$\Delta_{sub} S^{\circ}$ [b]	$\Delta_{vap} H^{\circ}$ [a]	$\Delta_f G^{\circ}_{ig}$ [a]	$\Delta_f H^{\circ}_{ig}$ [a]	S°_{ig} [b]
1,2-ethanediol	2	-333.3[d]	-461.8[e]	193[f]	54.60[g]	-32.9[h]	-72.9[i]		-6.87[j]				66.0[k]	-300.4[l]	-388.9[m]	311.84[n]
glycerol	3	-491.4[d]	-675.4[o]	240[f]	70.91[g]	-51.1[h]								-440.3[p]		
erythritol	4			310[f]	87.1[g]	-51.9[q]		-3.999[r]		47.9[s]	140[t]	309[u]				
threitol	4				86.13[v]											
xylitol	5	-791.87[w]	-1097.65[w]	346[f]	102.14[g]	-64.8[q]		4.0866[x]		68.9[s]	161[u]	309[u]				
d-arabitol	5			375[f]	102.6[v]											
l-arabitol	5			373[f]	102.6[v]											
ribitol	5			376[f]	102.7[v]											
iditol	6	-938.93[y]	-1309.94[y]													
mannitol	6	-942.2[d]		452[f]	119.71[g]	-97.5[q]		-0.1162[r]		97.3[s]	202[u]	351[u]				
sorbitol	6	-944.49[y]	-1312.62[y]	412[f]	119.16[g]	-78.5[z]								-866[aa]		
galactitol	6				118.6[v]											

[a] kJ mol^{-1}, [b] J mol^{-1} K^{-1}, [c] cm^3 mol^{-1}, [d] calculated from $\Delta_f G^{\circ}_{aq} = \Delta_h G^{\circ} + \Delta_f G^{\circ}_{ig}$, [e] calculated from $\Delta_f H^{\circ}_{aq} = \Delta_h H^{\circ} + \Delta_f H^{\circ}_{ig}$, [f] Lian *et al.* (1982), [g] DiPaola and Belleau (1977), [h] Plyasunov *et al.* (2006), [i] calculated from $\Delta_h H^{\circ} = \Delta_{sol} H^{\circ} - \Delta_{vap} H^{\circ}$, [j] Nichols *et al.* (1976), [k] Verevkin (2004), [l] calculated using $\Delta_f G^{\circ}_{ig}$ and $\Delta_f H^{\circ}_{ig}$ given in the table along with S° of the elements referenced from Cox *et al.* (1989) and the relation $\Delta_f G^{\circ}_{ig} = \Delta_f H^{\circ}_{ig} - T\Delta(S^{\circ}_{ig} - \sum S^{\circ}_{elements})$, [m] Verevkin *et al.* (2009), [n] Chao *et al.* (1986), [o] Bastos *et al.* (1988), [p] Verevkin *et al.* (2015), [q] calculated from $\Delta_h G^{\circ} = \Delta_{sol} G^{\circ} - \Delta_{sub} G^{\circ}$, [r] calculated using the relation $\Delta_{sol} G^{\circ} = -RT \ln K_{sp}$, where R is the gas constant, T is temperature, and K_{sp} is the solubility product equilibrium constant calculated from the number of grams of sugar alcohol soluble per 100 grams of water at room temperature reported by Dehn (1917), [s] calculated using $\Delta_f G^{\circ}_{sub}$ and $\Delta_f H^{\circ}_{sub}$ given in the table and the relation $\Delta_f G^{\circ}_{sub} = \Delta_f H^{\circ}_{sub} - T\Delta_{sub} S^{\circ}$), [t] Lopes Jesus *et al.* (2006), [u] Barone *et al.* (1990), [v] Chavez and Birch (1997), [w] da Silva *et al.* (2013), [x] Wang *et al.* (2013), [y] Tewari and Goldberg (1996), [z] calculated using values for $\Delta_f G^{\circ}_{aq}$ and $\Delta_f G^{\circ}_{ig}$ in the table and the relation $\Delta_h G^{\circ} = \Delta_f G^{\circ}_{aq} - \Delta_f G^{\circ}_{ig}$, [aa] Yaws (1997)

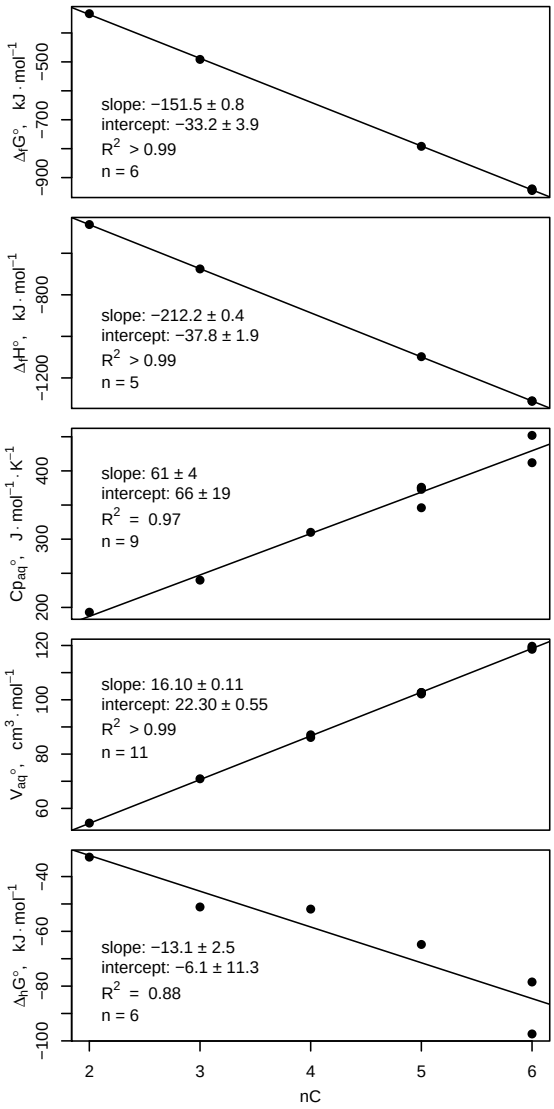


Figure 4.3: Linear regression of experimentally derived aqueous partial molal thermodynamic properties of sugar alcohols as a function of carbon length (nC). Numbers of regressed datapoints for each plot are displayed equal to 'n'. Literature sources for regressed data are given in Table 4.1.

Table 4.2: Partial molal thermodynamic properties of non-stereoisomeric sugar alcohols estimated from the linear regression of sugar alcohol diasteromer properties given in Table 4.1 with carbon length, as shown in Figure 4.3.

	nC	$\Delta_f G^{\circ}_{aq}$ [a]	$\Delta_f H^{\circ}_{aq}$ [a]	Cp°_{aq} [b]	V°_{aq} [c]	$\Delta_h G^{\circ}$ [a]
1,2,3,4-butanetetrol	4	-639.3	-886.6	308	86.7	-58.4
1,2,3,4,5-pentanepentol	5	-790.9	-1098.8	369	102.8	-71.5
1,2,3,4,5,6-hexanehexol	6	-942.4	-1311.0	429	118.9	-84.6

[a] kJ mol^{-1}, [b] J mol^{-1} K^{-1}, [c] cm^3 mol^{-1}

Table 4.3: Partial molal thermodynamic data for second order groups used in the estimation of BHP headgroup properties.

Group	Formula	$\Delta_f G^\circ_{aq}$ [a]	$\Delta_f H^\circ_{aq}$ [a]	S°_{aq} [b]	Cp°_{aq} [b]	V°_{aq} [c]	$\Delta_h G^\circ$ [a]	$\Delta_h H^\circ$ [a]	$\Delta_h S^\circ$ [b]	$\Delta_h Cp^\circ$ [b]	$\Delta_f H^\circ_{ig}$ [a]	S°_{ig} [b]	Cp°_{ig} [b]
C-(H)(O)(C)$_2$	CH	6.29[d]	-27.98[e]	-43.85[f]	26[g]	7.48[h]	-1.64[h]	-1.88[h]	-0.80[i]	6[h]	-26.10[j]	-43.05[j]	19.96[j]
C-(H)(C)$_3$	CH	34.07[d]	1.17[e]	-39.28[f]	3[g]	5.96[h]	-1.93[h]	2.34[h]	14.32[i]	-17[h]	-1.17[j]	-53.60[j]	20.08[j]
C-(H)$_2$(O)(C)	CH$_2$	-4.41[d]	-38.07[e]	23.51[f]	88[g]	17.25[h]	0.77[h]	-5.17[h]	-19.92[i]	68[h]	-32.90[j]	43.43[j]	20.33[j]
C-(H)$_2$(C)$_2$	CH$_2$	9.05[d]	-24.15[e]	25.07[f]	85[g]	15.61[h]	0.68[h]	-3.52[h]	-14.09[i]	62[h]	-20.63[j]	39.16[j]	22.89[j]
C-(H)$_3$(C)	CH$_3$	-16.35[d]	-50.45[e]	87.37[f]	158[g]	25.56[h]	3.72[h]	-8.19[h]	-39.95[i]	132[h]	-42.26[j]	127.32[j]	25.73[j]

[a] kJ mol^{-1}, [b] J mol^{-1} K^{-1}, [c] cm^3 mol^{-1}, [d] calculated using $\Delta_f H^\circ_{aq}$ and S°_{aq} given in the table along with S° of the elements referenced from Cox *et al.* (1989) and the relation $\Delta_f G^\circ_{aq} = \Delta_f H^\circ_{aq} - T\Delta(S^\circ_{aq} - \sum S^\circ_{elements})$, [e] calculated using $\Delta_h H^\circ$ and $\Delta_f H^\circ_{ig}$ given in the table and the relation $\Delta_f H^\circ_{aq} = \Delta_h H^\circ + \Delta_f H^\circ_{ig}$, [f] calculated using $\Delta_h S^\circ$ and S°_{ig} given in the table and the relation $S^\circ_{aq} = \Delta_h S^\circ + S^\circ_{ig}$, [g] calculated using $\Delta_h Cp^\circ$ and Cp°_{ig} given in the table and the relation $Cp^\circ_{aq} = \Delta_h Cp^\circ + Cp^\circ_{ig}$, [h] Plyasunov *et al.* (2004), [i] calculated using $\Delta_h G^\circ$ and $\Delta_h H^\circ$ given in the table and the relation $\Delta_h G^\circ = \Delta_h H^\circ - T\Delta_h S^\circ$, [j] Domalski and Hearing (1993)

in estimated properties for six-carbon-long BHP tetrol, pentol, and hexol headgroups (structure D). In step 4, the properties of 1,2-ethanediol (given in Table 4.1) are subtracted from structure D and those of ethanolamine (given in Table 4.4) are added, resulting in the estimated properties of BHP aminotriol, aminotetrol, and aminopentol headgroups (structure E). In step 5, the properties of ionized aminopolyol BHP headgroups (structure F) are obtained by subtracting the properties of ethanolamine and adding those of ethanolamminium given in Table 4.4.

4.2.4 Estimation of revised Helgeson-Kirkham-Flowers (HKF) equation of state parameters for BHP polar headgroups

Fitting parameters for the revised HKF equation of state were estimated for uncharged BHP polyol and aminopolyol headgroups using $\Delta_f G^\circ_{aq}$, $\Delta_f H^\circ_{aq}$, Cp°_{aq}, V°_{aq}, and $\Delta_h G^\circ_{aq}$ and the correlation strategies given in Plyasunov and Shock (2001) for aqueous nonelectrolytes. Charged ammoniumpolyol BHP headgroups HKF parameters were estimated with a hybrid scheme, where the correlation strategy in Sverjensky *et al.* (2014) was used to estimate a_1 for charged and uncharged aqueous species up to 60 kb, and the estimation methods in Shock and Helgeson (1990) were used to obtain ω (from S°_{aq}), c_1, c_2, a_2, and a_4. The parameter a_3 was solved from ω and the other a parameters using the volume-dependent portion of the HKF equation of state, Equation A-1 in Tanger and Helgeson (1988). Estimated thermodynamic properties and HKF equation of state parameters for aqueous BHP headgroups are reported in Table 4.5.

Table 4.4: Partial molal thermodynamic properties of aqueous chemical species used in the estimation of BHP aminopolyol headgroup properties.

	$\Delta_f G^\circ_{aq}$ [a]	$\Delta_f H^\circ_{aq}$ [a]	Cp°_{aq} [b]	V°_{aq} [c]	$\Delta_h G^\circ$ [a]
ethane	-16.29[d]	-103.21[d]			
ethanol	-181.29[d]	-287.2[d]			
methylamine					-19.1[e]
methylammonium					-315[f]
ethylamine	26.36[g]	-99.705[g]			
ethanolamine	-138.64[h]	-283.7[i]	176[j]	59.2[k]	-18.9[f]
ethanolammonium	-192.5[l]	-332.3[m]	180.9[n]	53.0[o]	-305[f]

[a] kJ mol^{-1}, [b] J mol^{-1} K^{-1}, [c] cm^3 mol^{-1}, [d] recommended value from the ORganic Compounds HYDration properties (ORCHYD) database compiled by Plyasunova *et al.* (2004), [e] Ben-Naim and Marcus (1984), [f] Pliego and Riveros (2000), [g] Shock and Helgeson (1990), [h] estimated from $\Delta_f G^\circ_{aq,ethanolamine} = \Delta_f G^\circ_{aq,ethylamine} + (\Delta_f G^\circ_{aq,ethanol} - \Delta_f G^\circ_{aq,ethane})$, [i] estimated from $\Delta_f H^\circ_{aq,ethanolamine} = \Delta_f H^\circ_{aq,ethylamine} + (\Delta_f H^\circ_{aq,ethanol} - \Delta_f H^\circ_{aq,ethane})$, [j] Cabani *et al.* (1981), [k] Maham *et al.* (1994), [l] calculated from $\Delta_f G^\circ_{aq,ethanolammonium}$ and $\Delta_f G^\circ_{aq,ethanolamine}$ in the table using the relation $\Delta_f G^\circ_{aq,ethanolammonium} = \Delta_f G^\circ_{aq,ethanolamine} - \Delta_r G^\circ_d$ where $\Delta_r G^\circ_d$ is the standard state Gibbs free energy of ethanolammonium dissociation taken as 53.90 kJ mol^{-1} from Hamborg and Versteeg (2009), [m] calculated from $\Delta_f H^\circ_{aq,ethanolammonium}$ and $\Delta_f H^\circ_{aq,ethanolamine}$ in the table using the relation $\Delta_f H^\circ_{aq,ethanolammonium} = \Delta_f H^\circ_{aq,ethanolamine} - \Delta_r H^\circ_d$ where $\Delta_r H^\circ_d$ is the standard state enthalpy change of ethanolammonium dissociation taken as 48.60 kJ mol^{-1} from Hamborg and Versteeg (2009), [n] calculated from $Cp^\circ_{aq,ethanolammonium}$ and $Cp^\circ_{aq,ethanolamine}$ in the table using the relation $Cp^\circ_{aq,ethanolammonium} = Cp^\circ_{aq,ethanolamine} - \Delta_r Cp^\circ_d$ where $\Delta_r Cp^\circ_d$ is the standard state isobaric heat capacity change of ethanolammonium dissociation taken as -4.9 J mol^{-1} K^{-1} from Bates and Pinching (1951), [o] calculated from $V^\circ_{aq,ethanolammonium}$ and $V^\circ_{aq,ethanolamine}$ in the table using the relation $V^\circ_{aq,ethanolammonium} = V^\circ_{aq,ethanolamine} - \Delta_r V^\circ_d$ where $\Delta_r V^\circ_d$ is the standard state volume change of ethanolammonium dissociation taken as 6.2 cm^3 mol^{-1} from Cabani *et al.* (1977a).

Table 4.5: Estimated thermodynamic standard state partial molal properties and Helgeson-Kirkham-Flowers equation of state parameters for aqueous BHP headgroups.

BHP headgroup	formula	$\Delta_f G^{\circ}_{aq}$ [a]	$\Delta_f H^{\circ}_{aq}$ [a]	S°_{aq} [b]	Cp°_{aq} [b]	V°_{aq} [c]	a_1 [d]$\times 10$	a_2 [e]$\times 10^{-2}$	a_3 [f]	a_4 [g]$\times 10^{-4}$	c_1 [b]	c_2 [f]$\times 10^{-4}$	ω [e]$\times 10^{-5}$	$\Delta_h G^{\circ}$ [a]
tetrol	$C_8H_{17}O_4$	-592.8	-974.1	15.38	576	139.66	129	51.8	99.4	-43.1	636	-27.6	0.424	-57.7
pentol	$C_8H_{17}O_5$	-753.4	-1162.1	25.91	552	140.15	134	42.7	85.3	-39.4	638	-39.2	0.610	-71.4
hexol	$C_8H_{17}O_6$	-914.0	-1350.2	36.43	528	140.63	138	33.5	70.9	-35.6	638	-50.9	0.766	-85.2
aminotriol	$C_8H_{18}NO_3$	-398.2	-796.0	18.23	559	144.26	130	63.2	117	-48.6	593	-15.7	0.196	-43.6
aminotetrol	$C_8H_{18}NO_4$	-558.8	-984.1	28.76	535	144.75	134	53.8	103	-44.6	595	-27.3	0.420	-57.4
aminopentol	$C_8H_{18}NO_5$	-719.4	-1172.1	39.28	511	145.23	138	44.4	87.6	-40.6	596	-39.0	0.606	-71.1
amminiumtriol	$C_8H_{19}NO_3^+$	-452.1	-844.6	36.01	564	138.06	198	-130	-188	32.6	1110	-259	1.84	-330
amminiumtetrol	$C_8H_{19}NO_4^+$	-612.7	-1033.0	46.53	540	138.55	202	-139	-204	36.8	1110	-271	1.86	-344
amminiumpentol	$C_8H_{19}NO_5^+$	-773.3	-1220.7	57.06	516	139.03	206	-149	-220	41.0	1110	-282	1.89	-358

These values were generated as described in the methods. [a] kJ mol^{-1}, [b] J mol^{-1} K^{-1}, [c] cm^3 mol^{-1}, [d] J mol^{-1} bar^{-1}, [e] J mol^{-1}, [f] J K mol^{-1} bar^{-1}, [g] J K mol^{-1}

4.2.5 Calculating BHP headgroup metastable equilibrium abundance and headgroup-predicted Eh

Methods used for calculating metastable equilibrium abundances for aqueous species as a function of Eh are described in Chapter 3 for hypothetical site-averaged free alkyl chains. All nine polyol and aminopolyol BHP headgroup structures were speciated as a function of Eh. Observed ratios of tetrafunctional and pentafunctional BHPs in samples were compared to thermodynamically predicted ratios. $\mathrm{Eh}_{BHP,head}$ was designated as the Eh where the relative ratio of tetrafunctional to pentafunctional BHP headgroups predicted by metastable equilibrium calculations matched the ratio observed in extracted BHPs from a sample. For example, an observed ratio of 86:14 tetrafunctional to pentafunctional BHPs was observed in the lipid extract of sample BP2, and using the temperature and basis species concentrations of BP2, this ratio was predicted to be most stable at an Eh of -0.164 volts; the headgroup-predicted Eh of BP2).

4.2.6 Calculating predicted fraction of aminopolyol BHP headgroups

A calculation was performed to predict the abundance of aminopolyol BHP headgroups (aminotriol, aminotetrol, and aminopentol) relative to headgroups without amine functional groups (tetrol, pentol, hexol). Calculations to predict relative abundance of BHP headgroups at the temperature and basis species concentrations were carried out as previously described. However, in reactions to form BHP headgroups from basis species, the concentration of electrons was set by headgroup-predicted Eh value for each sample rather than allowing it to be an independent variable. In this way, ratios of aminopolyols to polyols were predicted by thermodynamic calculations at Eh values that were themselves predicted from observed ratios of tetrafunctional

and pentafunctional BHP headgroups.

4.3 Results

4.3.1 Observed BHP headgroup functionality

Percent abundance of BHP headgroup functionality is shown for all hot spring samples in Figure 4.4. Tetrol (BHT with and without composite groups) was the dominant BHP headgroup in most samples, followed by pentols (BHpentol with and without composite groups). The ratio of pentol to tetrol exhibited a sharp increase in most samples representative of photosynthetic mats, such as BP4, BP5, and MS4. The mat sample of Octopus Spring, OS2, did not differ much from its upstream sample, OS1, however. Empress Pool did not have a visible cyanobacterial mat and still showed an increase in pentol to tetrol ratio within the temperature regime one would expect of a photosynthetic mat in the other three springs. Interestingly, multiple samples that were distal from their respective hot spring sources, such as BP6, MS5, and EP5, exhibited a significant decrease in their pentol to tetrol ratio. Samples BP6 and MS5 correspond to post-mat locations with microbial communities that are visually distinct from those in the mat. Sample EP5, while relatively indistinguishable from samples immediately upstream from it, had a temperature similar to MS5. BHPs with aminotriol headgroups were proportionally most abundant in samples belonging to Empress Pool, constituting about 10% of the total BHP headgroups observed at EP1 and EP4. Elsewhere, the abundance of aminotriol was rarely over 5%. Aminotetrol was not observed in this sample set. Hexafunctional headgroups were observed in low abundance only in photosynthetic mat samples; with 2-5% hexol in BP4 and MS4, and a trace amounts ($< 1\%$) of aminopentol in BP4.

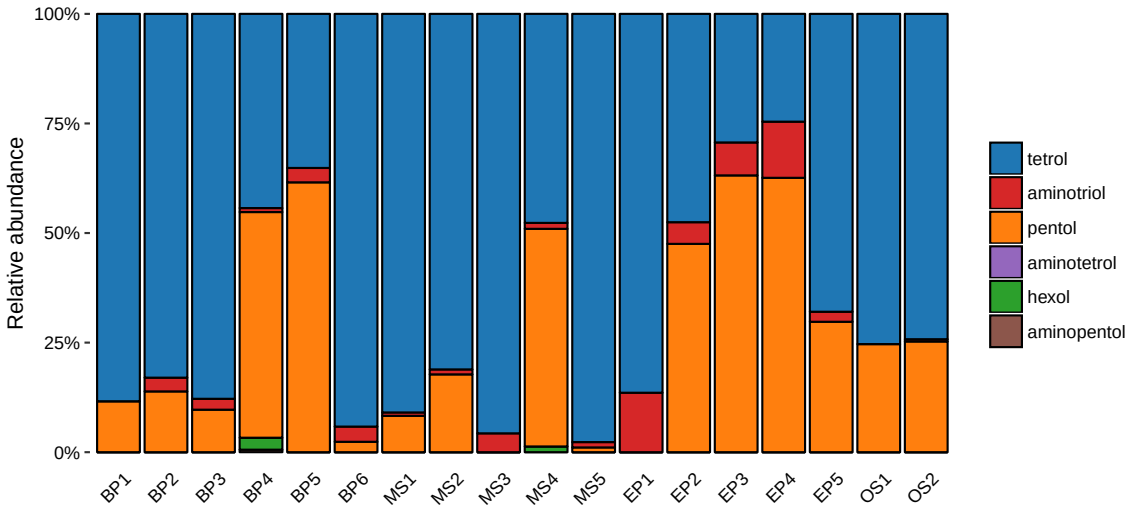


Figure 4.4: Relative abundances of BHP polar headgroup functionality observed in Yellowstone hot spring sediments and biofilms.

4.3.2 Z_C of BHP headgroups

The weighted Z_C of tetrafunctional, pentafunctional, and hexafunctional polyol and aminopolyol BHP headgroups (calculated for their non-composite form) are given in Table 4.6.

Each step increase in the functionality of a BHP headgroup is accompanied by a change of $+0.3\overline{6}$ in Z_C, and as such, samples with a greater ratio of higher-functionality BHP headgroups exhibited a higher weighted Z_C. As shown in Figure 4.5, weighted Z_C did not have a strong linear correlation with temperature ($R^2 = 0.20$), nor did it correlate with concentration of dissolved oxygen ($R^2 = 0.16$). Unlike the clear, linear increase of alkyl chain Z_C with decreasing temperature, the rise in BHP headgroup Z_C downstream is punctuated by a sudden drop, indicating the appearance of BHP headgroups with relatively reduced carbon in low-temperature samples. Recall that BHP headgroups tetrol and pentol comprised most of the headgroups in this sample set and that the pentol to tetrol ratio tended to be greatest in downstream mat samples spanning an approximate range of 40-74°C, resulting in the relatively high Z_C for BHP headgroups at these sites. The sudden decrease in pentol to tetrol ratio in the post-mat samples is responsible for the dip in Z_C at the lowest-temperature sites.

Taken together, Z_C as a predictive proxy for adaptation in BHP headgroup functionality temperature and chemical composition is not as easily interpretable as it appears to be for alkyl chains (Chapter 2). A more in-depth thermodynamic analysis was required to provide evidence that downstream alkyl chain distributions follow an energetically favorable trajectory with respect to both temperature chemical composition, including redox (Chapter 3). A similar approach was taken to predict whether observed distributions in BHP headgroup functionality are on an energetically favor-

Table 4.6: Select geochemical and physical data for samples, calculated and lipid-predicted redox potentials, and the average oxidation state of carbon in BHP headgroups.

				Weighted Z_C	Lipid-predicted Eh, volts	
Site	Sample	T, °C	pH	BHP headgroups	Eh_{alkyl}[a]	$Eh_{BHP,\ head}$[b]
Bison	BP1	89.0	7.23	-1.10	-0.504	-0.175
Pool	BP2	80.9	7.34	-1.09	-0.468	-0.164
	BP3	73.3	7.27	-1.09	-0.435	-0.153
	BP4	63.1	8.09	-0.78	-0.464	-0.153
	BP5	40.5	8.25	-0.42	-0.394	-0.110
	BP6	29.0	9.01	-1.09	-0.350	-0.190
Mound	MS1	91.0	8.81	-1.10	-0.658	-0.299
Spring	MS2	77.3	8.65	-1.07	-0.562	-0.245
	MS3	64.8	9.08	-1.13	-0.557	-0.292
	MS4	53.0	9.22	-0.65	-0.519	-0.210
	MS5	35.1	9.53	-1.12	-0.438	-0.251
Empress	EP1	82.2	5.78	-1.13	-0.383	-0.101
Pool	EP2	70.5	6.96	-0.83	-	-
	EP3	60.7	7.63	-0.51	-0.451	-0.113
	EP4	51.6	7.99	-0.60	-0.422	-0.120
	EP5	38.1	8.42	-0.75	-0.393	-0.141
Octopus	OS1	85.4	7.29	-1.00	-0.467	-0.159
Spring	OS2	59.8	8.27	-0.83	-0.477	-0.176

Dashed entries represent samples where a particular Eh could not be calculated or estimated. See text. [a] Eh predicted from alkyl chains of IPLs as reported in Chapter 3, [b] Eh predicted by BHP polar headgroups (this work).

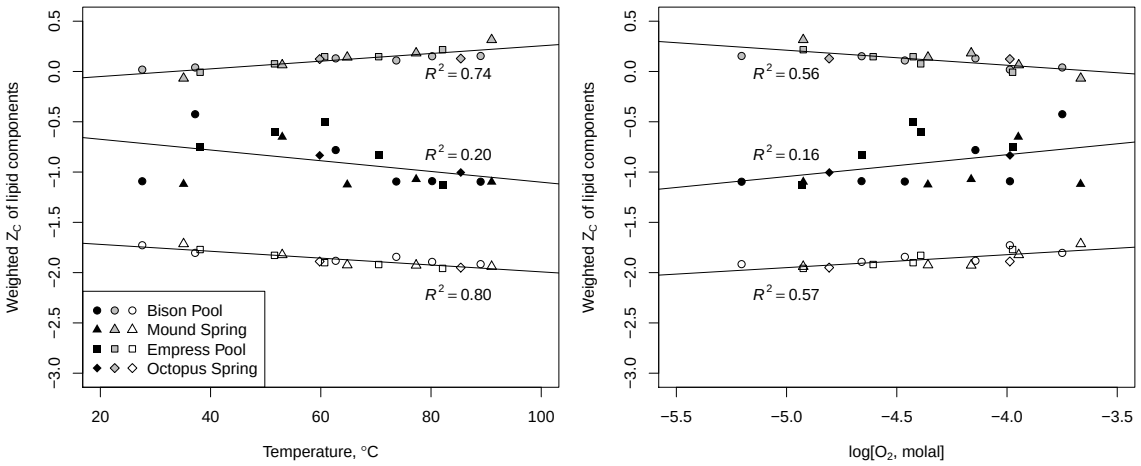
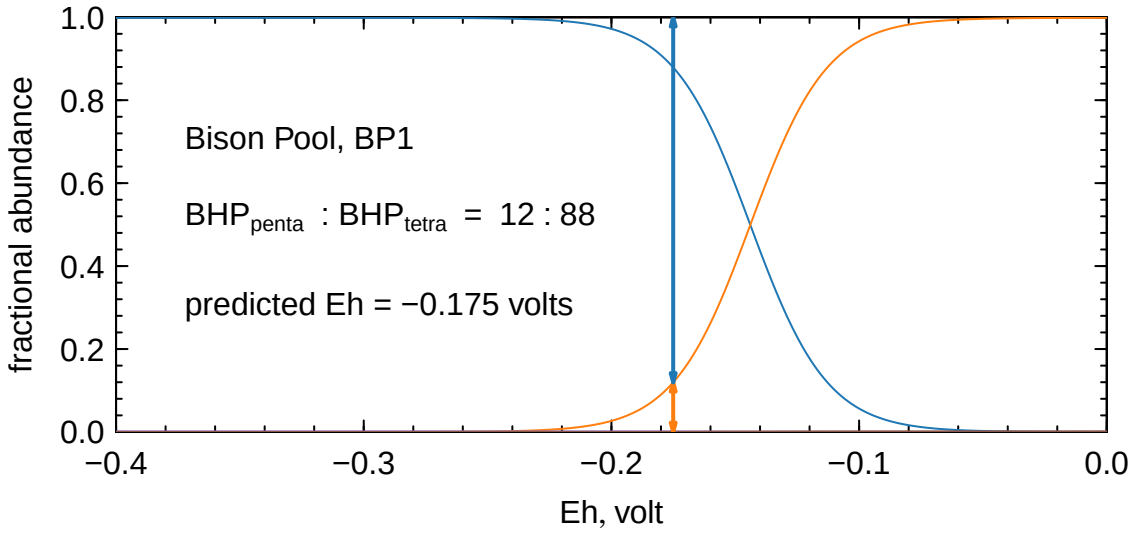


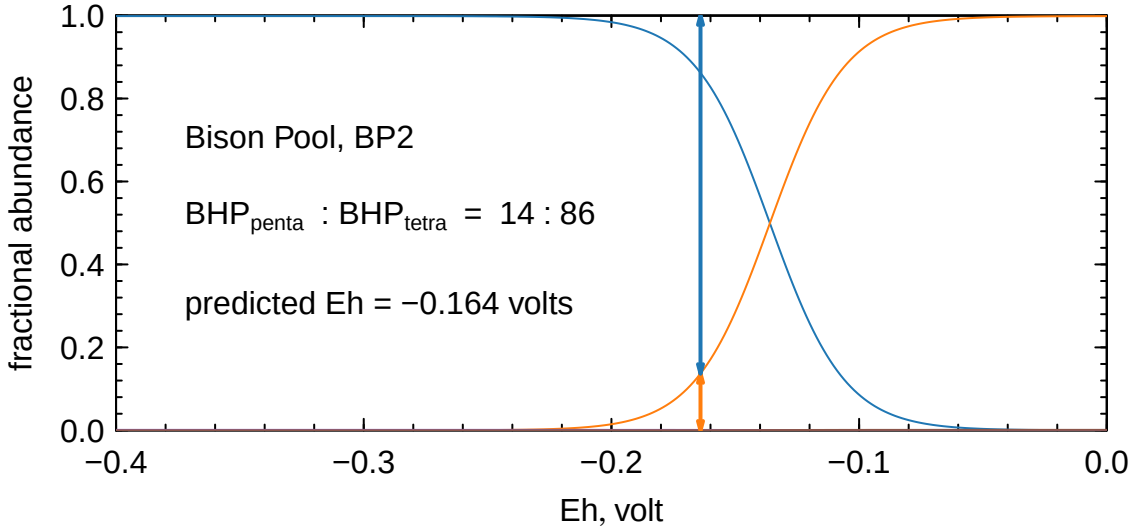
Figure 4.5: Weighted Z_C of thermophile lipid components as a function of temperature and dissolved oxygen concentration. Black symbols indicate BHP headgroups, and gray symbols indicate IPL headgroups. Empty symbols indicate IPL alkyl chains. Values for Z_C of IPL are from Chapter 2. Unweighted linear models were used to generate R^2 values.

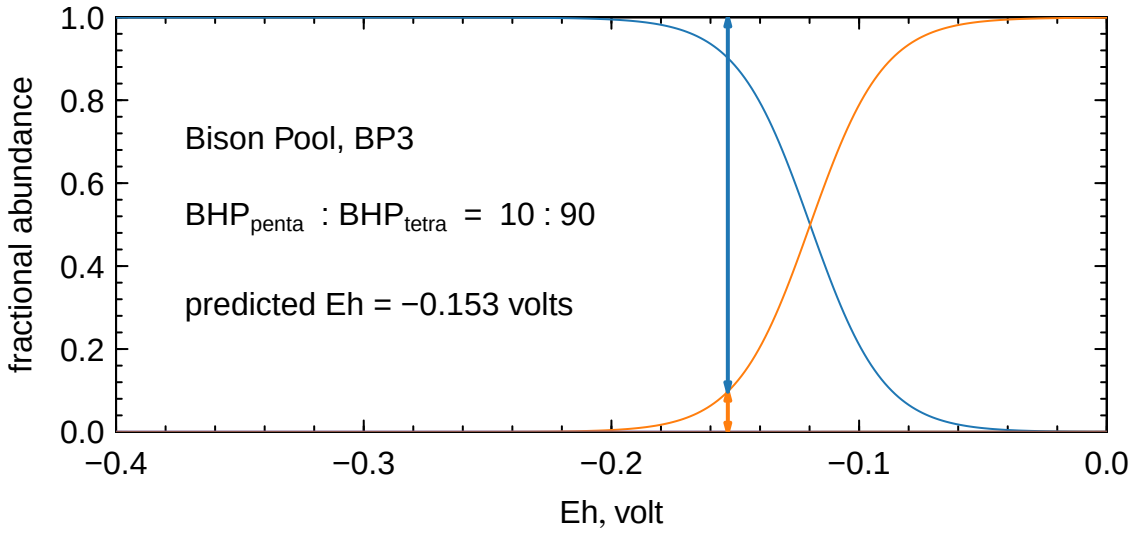
able trajectory downstream in such a way that is not apparent with Z_C.

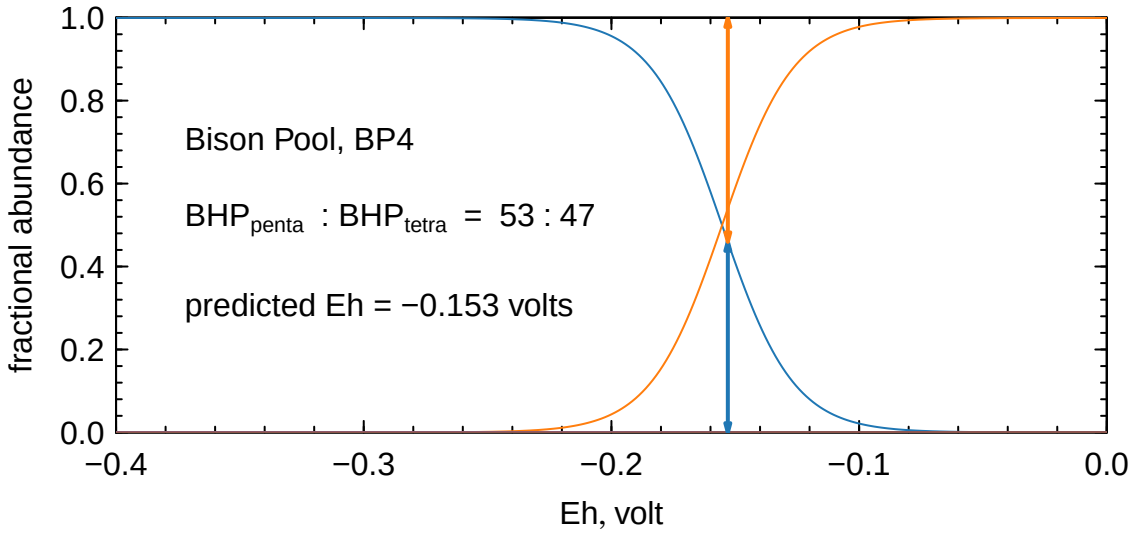
4.3.3 Thermodynamic predictions of $Eh_{BHP,\ head}$

Tetrafunctional BHP headgroups are predicted to be stable under increasingly reduced conditions relative to pentafunctional headgroups, which are stable under increasingly oxidized conditions, as shown in Figure 4.6 for Bison Pool samples. At the temperature and basis species concentrations of sample BP1, metastable equilibrium calculations predict a fractional abundance of nearly 1 for tetrafunctional BHP headgroups an Eh of -0.2 volts and below, while pentafunctional headgroups are predicted to take over at an Eh above -0.08 volts. At approximately -1.45 volts, pentafunctional and tetrafunctional BHP headgroups are predicted to have the equal abundance or a fractional abundance of 0.5 for each. Observed ratios of headgroup functionality in BHPs extracted from sample BP1 was about 12:88. The concentration of electrons in reactions to form BHP headgroups from basis species that would best predict a 12:88 penta- to tetrafunctionality ratio corresponds to an $Eh_{BHP,\ head}$ of -0.175 volts in sample BP1. In samples BP2 and BP3, similar ratios in functionality are observed (about 14:86 and 10:90, respectively), though increasingly oxidized conditions (-1.64 and -0.153 volts) are required to predict these ratios due to a decrease in temperature and a shift in chemical composition. Sample BP4 represents the onset of the photosynthetic microbial mat and marks the first time in Bison Pool where pentafunctional headgroups become more abundant than tetrafunctional headgroups by a ratio of about 53:47. Intriguingly, metastable equilibrium predicts that at the temperature and basis species concentrations of BP4, the $Eh_{BHP,\ head}$ that best describes this headgroup functionality ratio is -0.153 volts, which is identical to that of BP3. An $Eh_{BHP,\ head}$ of -0.110, the most oxidized conditions predicted for Bison Pool yet, is needed to explain the pentafunctional to tetrafunctional headgroup ra-


1.0
0.8
0.6
0.4
0.2
0.0
fractional abundance
Bison Pool, BP1
BHP$_{penta}$: BHP$_{tetra}$ = 12 : 88
predicted Eh = −0.175 volts
−0.4
−0.3
−0.2
−0.1
0.0
Eh, volt


1.0
0.8
0.6
0.4
0.2
0.0
fractional abundance
Bison Pool, BP2
BHP$_{penta}$: BHP$_{tetra}$ = 14 : 86
predicted Eh = −0.164 volts
−0.4
−0.3
−0.2
−0.1
0.0
Eh, volt


1.0
0.8
0.6
0.4
0.2
0.0
fractional abundance
Bison Pool, BP3
BHP$_{penta}$: BHP$_{tetra}$ = 10 : 90
predicted Eh = −0.153 volts
−0.4
−0.3
−0.2
−0.1
0.0
Eh, volt


1.0
0.8
0.6
0.4
0.2
0.0
fractional abundance
Bison Pool, BP4
BHP$_{penta}$: BHP$_{tetra}$ = 53 : 47
predicted Eh = −0.153 volts
−0.4
−0.3
−0.2
−0.1
0.0
Eh, volt

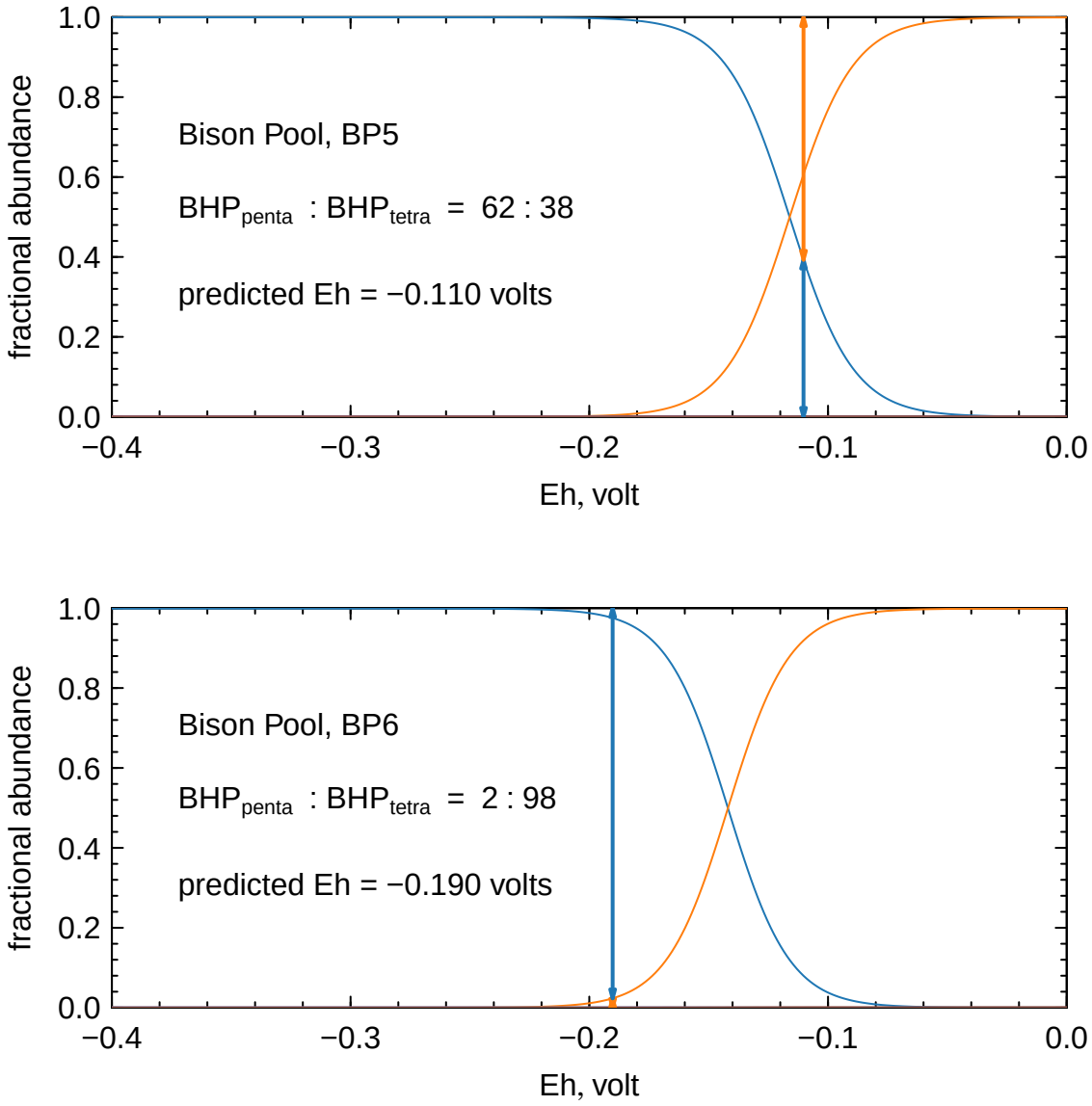


Figure 4.6: Predicted metastable equilibrium abundance of tetrafunctional (blue) and pentafunctional (orange) BHP headgroups in Bison Pool samples compared to observed ratios (blue to orange vertical arrow lengths) with corresponding BHP headgroup-predicted Eh indicated for each.

tio at BP5, which reached its peak of 62:38 for this spring. In the post-mat sample BP6, the abundance of pentafunctional headgroups had diminished substantially to a ratio of 2:98 predicted by an $Eh_{BHP,\ head}$ of -0.190 volts. Metastable equilibrium abundances of pentafunctional and tetrafunctional BHP headgroups as a function of Eh in Mound Spring, Empress Pool, and Octopus Spring samples are shown in Table C.1. $Eh_{BHP,\ head}$ values for all samples are given in Table 4.6.

Temperature was not found to correlate with $Eh_{BHP,\ head}$ ($R^2 = 0.02$), as shown in Figure 4.7. Compared to the weak correlation of Eh predicted by Eh_{alkyl}, this may suggest BHP structural adaptation may be less constrained by temperature than adaptations in alkyl chain structure. However, a strong positive correlation was found between $Eh_{BHP,\ head}$ and sample Eh calculated from various redox couples reported in Table 3.9 of Chapter 3; NO_2^-/NO_3^- ($R^2 = 0.74$), NH_4^+/NO_3^- ($R^2 = 0.79$), and HS^-/SO_4^{2-} ($R^2 = 0.75$). The strongest correlation was found between $Eh_{BHP,\ head}$ and Eh calculated from O_2/H_2O ($R^2 = 0.84$, Figure 4.7). The slope of the trend was 1.06±0.12, suggesting a 1:1 relationship between Eh_{O_2/H_2O} and $Eh_{BHP,\ head}$. The regression of $Eh_{BHP,\ head}$ as a function of Eh_{O_2/H_2O} runs approximately parallel to that of Eh_{alkyl} but predicts an Eh that is about 0.40±0.14 volts more oxidized based on the difference in y-intercepts.

4.3.4 Thermodynamic predictions of aminopolyol abundance

As shown in the left panel of Figure 4.8, the fraction of aminoBHPs in observed in lipid extracts is approximately 1-2 orders of magnitude greater than what is predicted by thermodynamic predictions of metastable equilibrium. This could be due to the sources of error in the way the thermodynamic properties of BHP headgroups were estimated, an unrepresentative choice of basis species, uncertainties in the underlying experimental data, *etc.*. It is possible that aminopolyol headgroup distributions

in nature are responding to variables that are not captured in this thermodynamic model or its underlying energetic assumptions. If one assumes the reverse is true, the discrepancy between predicted and observed fractions could be explained potential differences in ionization efficiency between BHPs with aminopolyol and polyol headgroups during LC-MS analysis. Due to a lack of commercially available standards, response factor correction could not be performed during the quantification of BHP structures. For the fraction of predicted aminopolyol headgroups to approach observation, BHP LC-MS integrated peak areas would need a response factor correction of about 0.04x or 25x for non-aminopolyols (Figure 4.8, right panel). This is not outside the realm of possibility; Cooke *et al.* (2008) and van Winden *et al.* (2012) reported that peak areas of BHPs with nitrogen were approximately 12/8 as intense as those without among a suite of five BHP standards. That roughly translates to a response correction of about 8/12x (0.7x) to nitrogen-bearing BHP peak areas to bring them into approximate consistency with BHPs without nitrogen. It should be noted that these response factors were obtained using a different set of instrumentation than that used in this study, so reported response factors may not represent peak area correction values relevant to this analysis. That said, general differences in response factors have been observed between nitrogenous and non-nitrogenous authentic lipid standards on the same LC-MS model used in this study; though these standards were non-BHP intact polar lipids. As described in Chapter 2, many of the response factor assignments to lipids in samples without a representative headgroup standard were guided by the presence or absence of nitrogen in their chemical structure, as peak area intensity of standards with nitrogen tended to be greater than those without. This is not proof that BHP peak areas would behave in this way, since relative peak area intensity of analytes could be greatly affected by the LC method and mass spectrometer ionization source (which are not the same between the analysis of IPLs in

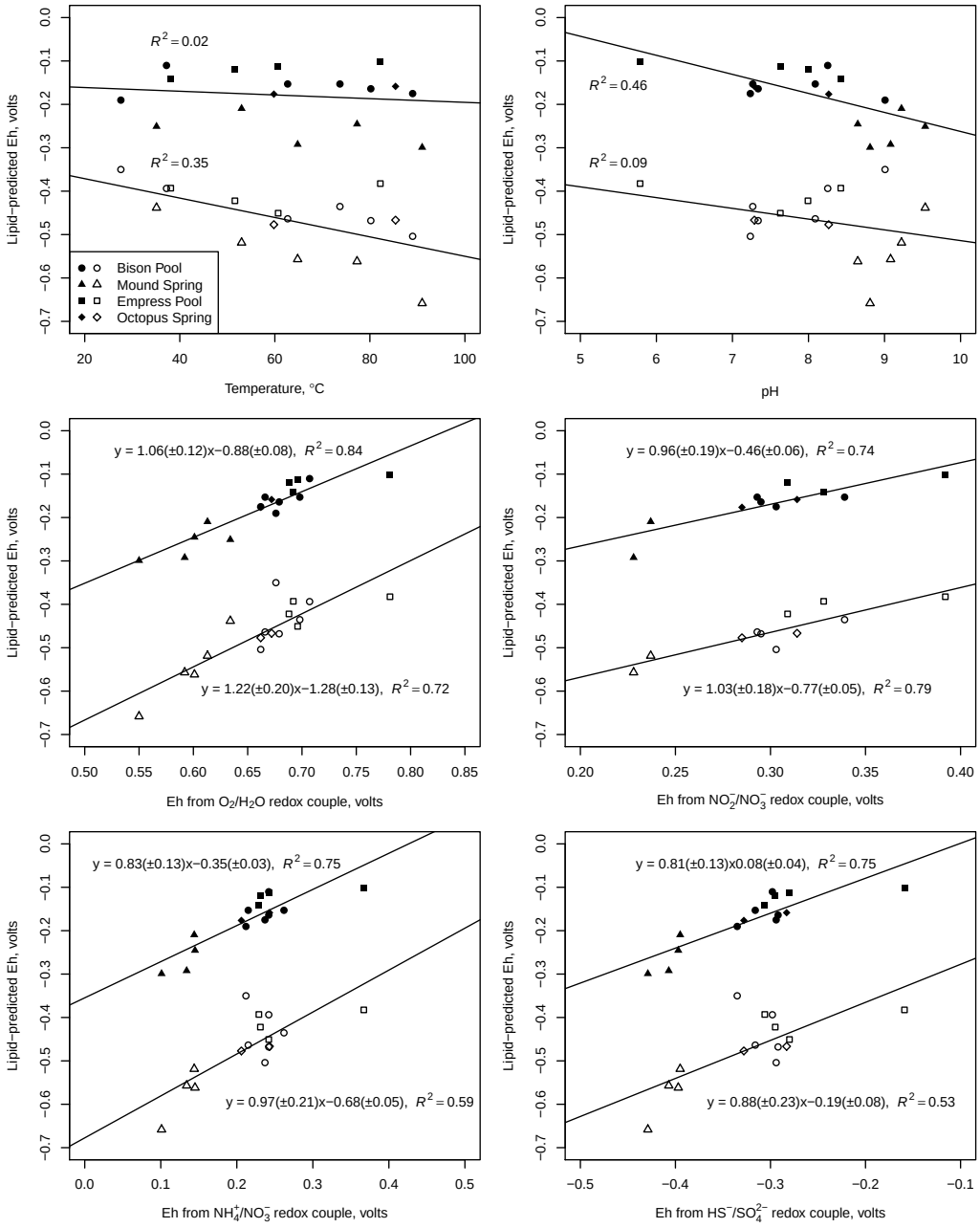


Figure 4.7: Lipid-predicted Eh as a function of temperature and redox potential. Black symbols indicate BHP headgroups, and gray symbols indicate IPL headgroups. Empty symbols indicate site-averaged free IPL alkyl chains. Values for alkyl chain-predicted Eh are from Chapter 3. Standard errors of slopes and intercepts of linear fits are shown in parentheses and do not include propagated uncertainties from underlying thermodynamic estimates. Unweighted linear models were used to generate R^2 values.

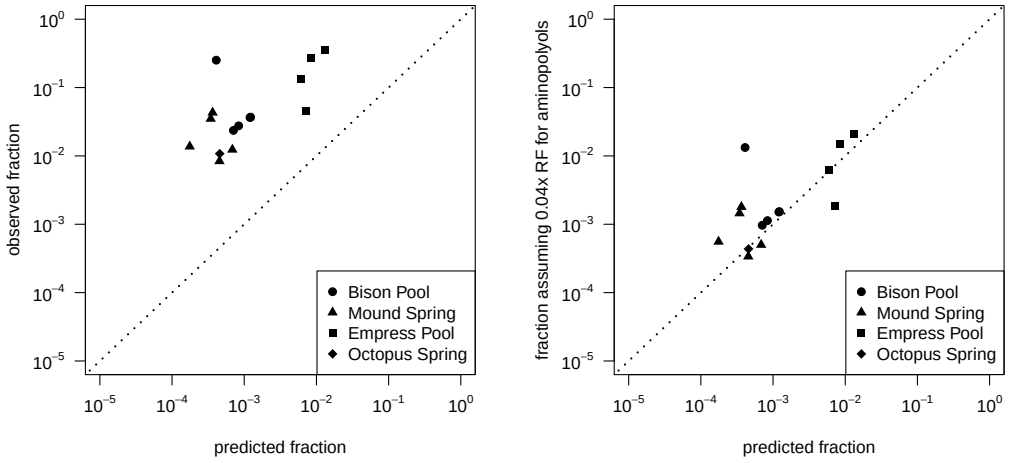


Figure 4.8: Predicted and observed fraction of aminopolyol BHP headgroups without response factor correction (left) and with a hypothetical response factor of 0.04x for aminopolyols relative to nonaminopolyols (right). Samples BP1 and OS2 are not plotted on this logarithmic scale because their observed fraction of aminopolyols was zero. ES2 is not plotted because a water sample was not collected, therefore, a predicted abundance based on concentrations of basis species could not be performed for this sample.

Chapter 2 and BHPs in this study); though it remains plausible that aminopolyol and polyol BHPs may have different response factors.

4.4 Discussion

The thermodynamic predictions of BHP headgroup favorability along Eh gradients performed in this work were carried out in a fashion similar to the estimations performed on site-averaged alkyl chains in Chapter 3. However, rather than calculating thermodynamic properties of hypothetical site-average lipid structures, this study aimed at estimating the thermodynamic properties of observed aqueous BHP polar headgroup structures reported in the literature and then using these properties to calculate equilibrium abundances of BHP headgroups along Eh gradients. Each Eh value that predicts an equilibrium-abundance BHP ratio that matched ratios observed in a sample lipid extract was denoted the 'BHP headgroup-predicted Eh' for that sample. These values of $Eh_{BHP,\ head}$ were compared to the Eh values calculated for samples from inorganic species reported in Chapter 3 in to determine whether there was a strong positive correlation. Such a correlation would suggest the abundances of BHP headgroups are following an energetically favorable trajectory in the context of hot spring temperature and chemical composition. The contribution of the BHP hydrophobic region was not included in thermodynamic calculations because their structure is relatively conserved among BHPs and was not expected to affect predicted BHP headgroup distributions significantly.

The results in $Eh_{BHP,\ head}$ can perhaps be best discussed in the context of a more familiar example; carbonate speciation with pH. As shown in Figure 4.9A, aqueous carbonate species chemically speciate at room temperature as a function of pH according to thermodynamic equilibrium. Dissolved carbon dioxide (CO_2) is the dominant species below about pH 6.2, bicarbonate (HCO_3^-) between 6.2 and 10.3, and carbon-

ate (CO_3^{2-}) above 10.3. If one did not know the pH of the system but could measure the concentration of carbonate species, the fractional abundance of carbonate could be used to solve for pH. For instance, a measured ratio of 75:25 CO_3^{2-} to HCO_3^- would indicate a pH of about 10.8 at room temperature or a pH of about 10.6 at 90°C (Figure 4.9B). Fractional abundance at 90°C was calculated from equilibrium constants predicted with the revised Helgeson-Kirkham-Flowers (HKF) equation of state (Shock *et al.*, 1992) using thermodynamic data and parameters reported in Shock and Helgeson (1988) and Plyasunov and Shock (2001) for aqueous carbonate species.

Reaction kinetics are rapid in the carbonate system, so a shift in pH results in a near-immediate re-equilibration of carbonate species into the fractional abundance at the new pH. In a hypothetical system where microbial membrane lipids are allowed to interchange rapidly in response to the chemical and thermal context of the surroundings, similar observations might be expected with respect to the fractional abundance of BHP headgroups. As an example, an increase in ammonia concentration might favor an increase in the fractional abundance of nitrogen-bearing aminopolyols. Likewise, a shift from relatively reduced to oxidized redox potential might favor the fractional abundance of BHP headgroups with a greater number of oxidized functional groups. As with the carbonate system, BHP headgroup fractionation would then represent the most thermodynamically stable equilibrium abundances based on the temperature and chemical conditions of the surroundings. By measuring the fractional abundance of BHP headgroups, it would then be possible to solve for a single chemical unknown assuming all other relevant chemical variables are known, much like predicting pH from the fraction of aqueous carbonate species. Furthermore, if thermodynamic properties were available for aqueous BHP headgroups, the HKF equation of state could be used to solve for a chemical unknown at elevated

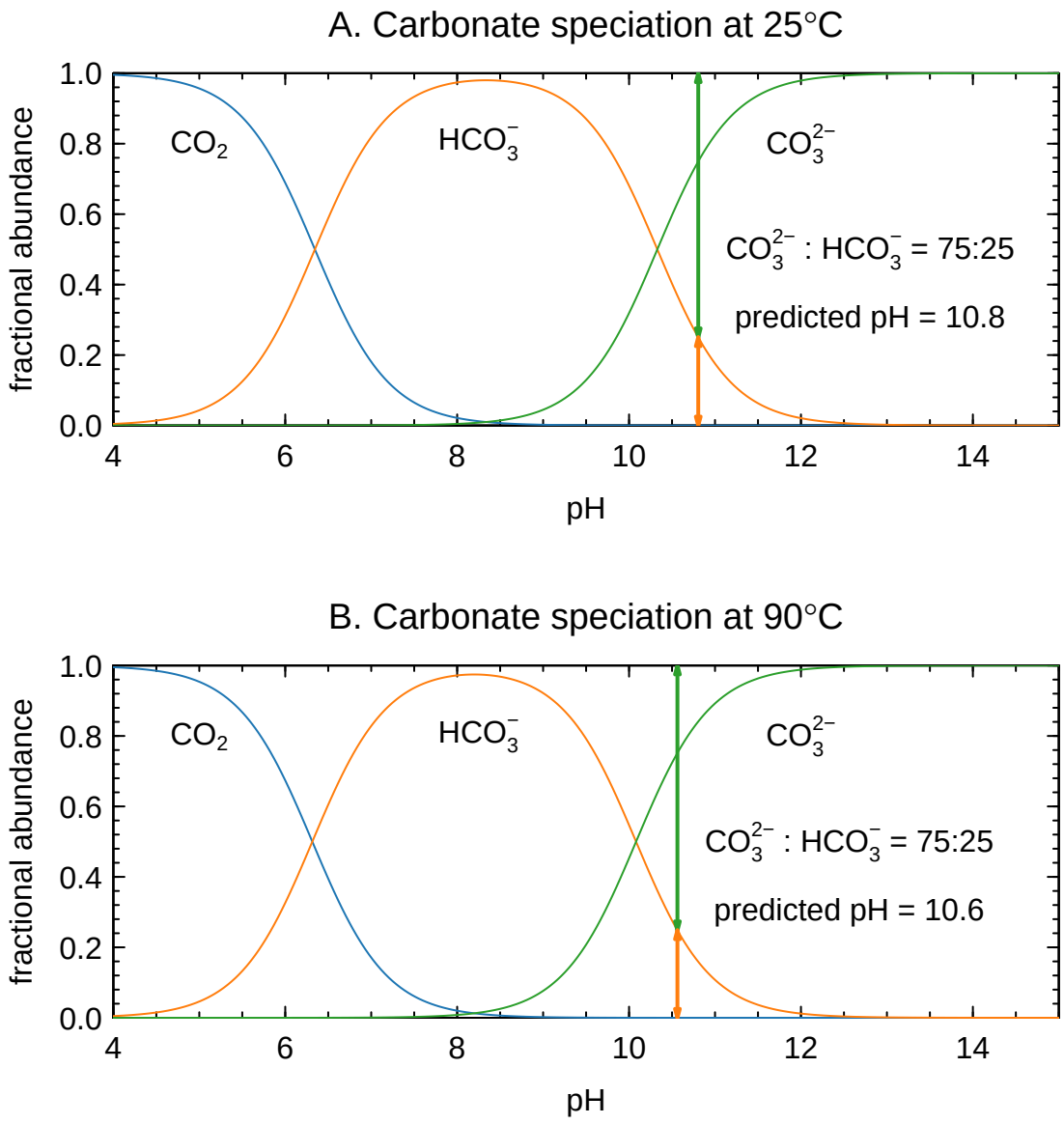


Figure 4.9: Predicted fractional abundance of aqueous carbonate species as a function of pH at 25°C (A) and 90°C (B). Vertical arrows show the pH predicted by a ratio of 75:25 CO_3^{-2} to HCO_3^-.

temperatures, such as those found in hot springs.

Based on what is known about BHP biosynthetic pathways in bacteria, however, it is unlikely that a change in chemical composition would result in reconfiguration of BHP headgroups at a rate even remotely comparable to that of the carbonate system. The BHP formylhopane is thought to be a precursor molecule that can either be reduced to a tetrol headgroup, or aminated to an aminotriol headgroup by the aminotransferase HpnO (Welander *et al.*, 2012). Therefore, it is unlikely that there would be a direct reaction of a polyol headgroup with ammonia to form an aminopolyol headgroup, as was posited earlier. This does not preclude the possibility that thermodynamic equilibrium among BHP headgroups could be converged upon indirectly; increased ammonia concentration might drive the glutamate dehydrogenase-catalyzed reaction of α-ketoglutarate to form glutamate (Lightfoot *et al.*, 1988), followed by the HpnO-catalyzed reaction of glutamate and formylhopane to α-ketoglutarate and a BHP aminopolyol, conceivably leading to an increase in the fractional abundance of aminopolyol BHP headgroups. However, experiments performed on *Desulfovibrio bastinii* showed no obvious difference in aminopolyol headgroup fraction between cultures growing with and without ammonium (Blumenberg *et al.*, 2012), though it should be noted the studied microorganism was capable of fixing its own nitrogen. It is still uncertain, then, whether bacteria metabolically shift fractional abundances of their BHP headgroups into configurations approaching thermodynamic equilibrium (at least in the case of ammonia and aminopolyols).

Whether or not bacteria shift their BHP headgroup distributions in such a way that approaches thermodynamic equilibrium within short, observable timescales, thermodynamically favorable assemblages of BHP headgroups may still be converged upon over the course of evolution. Life is not a system at equilibrium, so thermodynamic predictions of biomolecular compositions may only be expected to go so far. How-

ever, life must still obey the laws of thermodynamics and must construct itself out of the nutrients and energy supplies available to it. Organisms that make the best use of their surroundings have a competitive edge and are said to be 'adapted'. It is conceivable that organisms would save cellular energy over the course of evolution by adapting functional biomolecules that are more cost-effective in the temperature and chemical composition of their surroundings, or in other words, minimizing the Gibbs free energy of reactions to form biomolecules from available chemical species. These Gibbs free energies may never reach 0, but trends suggesting minimization could indicate an evolutionary approach toward equilibrium such that energy is saved while remaining competitive.

Equilibrium assumptions were used to evaluate observations of BHP headgroup distributions as a function of temperature and chemical composition. Thermodynamically predicted Eh values had a strong positive correlation with Eh independently calculated from measured concentrations of redox-sensitive inorganic species. These trends suggest BHP headgroup compositions have adapted to remain functional while minimizing Gibbs free energy of formation from available chemical species. In sum, observed abundances of BHP headgroups in this work may represent energetically driven distributions to provide functional membranes along natural thermal and chemical gradients.

4.5 Conclusion

The abundance-weighted Z_C of BHP headgroups does not have a clear correlation with temperature or dissolved oxygen concentration. However, Eh values calculated from inorganic redox couples are highly correlated with Eh values predicted by thermodynamic equilibrium to result in ratios of pentafunctional and tetrafunctional BHP headgroups observed in thermophile lipids. This provides evidence that observed

BHP distributions in microbial communities are approaching to equilibrium with respect to their biomolecular composition and supports the hypothesis that adaptations in lipid headgroups maximize function while minimizing cost. This conclusion was only reached after performing a thermodynamic analysis and could not be predicted from Z_C of BHP headgroups alone. This suggests that trends in Z_C for IPL headgroups reported in 2 are not necessarily energetically cost-ineffective. As with BHP headgroups, a thermodynamic analysis of IPL headgroups may reveal energetically favorable downstream trends.

Chapter 5

CONCLUSION

5.1 Summary

Organisms require functioning lipids across the entire range of conditions where life is found. As a result, there is an enormous variety of lipid adaptations in biology. As was discussed in this dissertation, however, certain lipid structures are known to function effectively in very different sets of environmental conditions; *e.g.*, glycerol dialkyl glycerol tetraethers (GDGTs) can create stable lipid membranes in both thermophiles and mesophiles (Schouten *et al.*, 2000). Likewise, very different sets of lipid distributions may function effectively in the similar environmental conditions; *e.g.*, GDGTs or non-GDGTs can create stable lipid membranes in mesophiles. Any given lipid adaptation, whether it is a double bond, a longer alkyl chain, or a particular headgroup, will have an energetic cost associated with its biosynthesis. These costs are predicted to change depending on the temperature, pressure, and chemical context of the surroundings.

This work provided evidence to support the hypothesis that natural selection favors convergence upon lipid compositions that maximize function while also minimizing energetic cost. First, the abundance-weighted Z_C of lipids sampled from microorganisms from four alkaline hot springs was shown to correlate positively with the oxidation state of the environment. Adaptations to provide alkyl chain functionality and membrane fluidity were themselves oxidized under oxidized conditions and reduced under reduced conditions, which I hypothesized was energetically cost-efficient. A subsequent thermodynamic analysis provided evidence to support this hypothesis

for intact polar lipid (IPL) alkyl chains by predicting that observed IPL distributions were stable relative to each other along redox gradients. Additionally, values of Eh predicted to result in favorable alkyl chain distributions were predicted to be correlated positively with environmental Eh. Lipid headgroups were also investigated and yielded results similar to those of alkyl chains; that hot spring Eh was correlated positively with Eh predicted by energetically stable configurations of observed headgroup ratios. In sum, this work suggests structural adaptations in lipid distributions are not only present because they function for their source organisms, but because they are cost-effective to produce within the temperature and geochemical context of their surroundings.

The thermodynamic studies described in this dissertation did not include calculations involving individual reactions of biosynthesis pathways or intracellular concentrations of chemical species in their evaluation of lipid cost. Instead, net reactions to form lipids from measured concentrations of bioavailable solutes alone provided evidence that observed thermophile lipid distributions may be energetically cost-effective. It is essential for cells to regulate intracellular conditions such as pH and solute concentrations for metabolic pathways to proceed. Maintaining chemical disequilibrium from the external environment requires energy. However, I would argue this energy requirement is based on how well a cell is adapted to its surroundings; energetic shortcuts to synthesize necessary biomolecules from available materials in the surroundings would offset a portion of this energy requirement. This raises the possibility that conditions *outside* of the cell matter more for the energetic costs of biomolecules than the conditions *inside* the cell. This concept could serve as a foundation for studies in molecular biology to evaluate potential costs of biomolecules synthesized by microbial communities in a variety of natural systems and in culture.

5.2 Future Work

Studies of lipid distributions in natural systems could be mainly exploratory (*i.e.*, which lipids are found under which sets of geochemical conditions) but could be supplemented by calculation of Z_C or thermodynamic predictions to allow quantitative assessment of lipid distributions in terms of energetic cost. These studies could also consider the phylogeny of source organisms, which was not done here. For instance, would a cyanobacteria-dominated microbial community that is adapted to thrive in water rich in hydrogen and methane, such as conditions found in fluids discharged from serpentinizing systems (McCollom and Bach, 2009), tend to produce lipid compositions with reduced carbon relative to cyanobacteria-dominated communities in oxidized ocean surface water? How phylogenetically related can two microbial communities be while having drastically different lipid compositions, and how might functional necessity and energetic cost help explain these differences?

These types of questions could also be probed in laboratory experiments. Various growth media of known chemical compositions could be inoculated with a mixed culture to isolate sub-sets of microbes that grow best under those conditions. Thermodynamic calculations based on those described in this work might then be used to predict whether the lipid compositions of isolated sub-sets tend to be energetically favorable in the chemical conditions of their own growth media relative to all others in the experiment. Would thermodynamic predictions offer greater insight into the expression of observed lipid distributions in these various treatments than the phylogenetic compositions of the isolated microbial sub-sets?

Future studies should attempt to be as quantitative as possible when measuring lipid abundances to allow estimation of weighted molecular properties like the average oxidation state of carbon (Z_C) or partial molal thermodynamic properties such as

Gibbs free energies of formation, volumes, entropies, and the like. The power to describe lipid distributions with of Z_C or thermodynamic calculations depends on the physical and chemical parameters of the system. As such, I recommend that these studies strive to be as complete as possible when measuring temperature, pressure, and chemical composition of the natural system or growth media.

Thermodynamic calculations could explore the formation of lipids from a different set of basis species than the 'autotrophic' set used in this work. It may be more appropriate to use acetate, for instance, to predict lipid distributions in a predominantly heterotrophic microbial community. Future studies could focus on reducing sources of error to aid in the comparison of thermodynamically predicted and analytically observed lipid abundances. Constraining sources of error introduced by lack of authentic response factors may offer a better estimation of aminopolyol BHP abundance. Alternate estimation schemes could be explored to estimate lipid thermodynamic properties. Liquid-phase or crystalline-phase alkyl chains, or a combination of the two, could be included in formation reactions from aqueous basis species to simulate the formation of a membrane lipid; this may offer a better predictive model than the formation of free aqueous lipids. Alternative methods for predicting complex distributions of lipids from basis species could be explored, and results could be compared to outcomes produced by the hypothetical site-averaged chains.

5.3 Broader Impacts

Interpreting the expression of lipids in terms of functional benefit and energy cost has the potential to change our understanding of the evolution of lipids, geochemical controls on lipid distributions, and ultimately how biomarkers are interpreted in the rock record.

Biochemistry emerged from geochemical processes, though it is still unclear when

lipids emerged in the history of life. Some theories posit that lipids were essential for metabolic encapsulation, growth, and division in the first proto-organisms (Deamer *et al.*, 2002; Luisi, 2016). Placing abiotic lipid synthesis in a thermodynamic framework could help narrow down geochemical conditions conducive to spontaneous lipid formation that could then be reconciled with the conditions proposed to allow early growth and division. It has also been theorized that lipid membranes evolved later. These theories hold that the last universal common ancestor (LUCA) that led to modern Bacteria, Archaea, and Eukarya lacked lipid membranes (Koga *et al.*, 1998; Martin and Russell, 2003), or had membranes with simpler lipid compositions (Sojo *et al.*, 2014). The notion that biology maximizes lipid function and minimizes energy cost could be applied to aid understanding of evolutionary pressures imposed by geochemistry that led to the divergence of lipid synthesis pathways among prokaryotes and eukaryotes after LUCA. While these are meant to represent specific examples where lipids are the focus, I am optimistic that thermodynamic assessments of the energetic costs of biomolecules would benefit evolutionary studies in general.

Lipid biomarkers have great potential to reveal detail about the environmental conditions experienced by their source communities, as well as geochemistry experienced by sedimentary rocks after diagenesis. In Ourisson *et al.* (1979), the authors report complex distributions of hopanoid lipid degradation products from a variety of sediments. They reported that hundreds of chemical structures were observed in these biomarkers, ranging from carboxylic acids to aldehydes to hydrocarbons to aromatic compounds, and emphasized that a wealth of information about geologic history could be unlocked if more could be learned about geochemical influences on the transformation of biomarkers. A review by Ourisson and Albrecht (1992) included speculation about patterns in hopanoid degradation products expected to form under oxidized and reduced conditions, but without a thermodynamic framework, these patterns can

only be interpreted qualitatively. This dissertation focused on thermal and geochemical controls on lipid expression in active, living systems, but thermodynamic data could be expanded to include biomarker degradation products. Thermodynamic calculations could potentially revolutionize the interpretation of biomarkers by allowing quantitative predictions regarding the chain of geologic events that led to observed biomarker distributions in the rock record.

5.4 Concluding Remarks

Geochemistry dictates what is necessary for habitability. This dissertation explored what that means for lipid adaptation; a lipid composition must function within the conditions experienced by the microbial community that produced it, and it must also have an acceptable energetic cost. This energetic cost is dictated by the temperature, pressure, and chemical composition of the system. A microbial community adapted to its surroundings is one that has evolved lipid distributions that are both functional and cost-effective. The composition of lipid structures produced by a microbial community should therefore reflect this interplay of energetic cost and functional benefit. Energetic cost can be assessed quantitatively, but this requires a thermodynamic framework to work within. This dissertation provides a step in that direction. Future work aimed at developing this framework further has the potential to shed great insight into why life produces the chemical structures that it does; not only with respect to lipid adaptation in thermophiles but for biomolecular adaptation across all of biology.

Amenabar, M. J., E. L. Shock, E. E. Roden, J. W. Peters and E. S. Boyd, "Microbial substrate preference dictated by energy demand rather than supply", Nature Geoscience **10**, 8, 577–581 (2017).

Amend, J. and E. Shock, "Energetics of amino acid synthesis in hydrothermal ecosystems", Science **281**, 5383, 1659–1662 (1998).

Barone, G., G. Della Gatta, D. Ferro and V. Piacente, "Enthalpies and entropies of sublimation, vaporization and fusion of nine polyhydric alcohols", Journal of the Chemical Society, Faraday Transactions **86**, 1, 75–79 (1990).

Bastos, M., S.-O. Nilsson, M. D. R. da Silva, M. A. R. da Silva and I. Wadsö, "Thermodynamic properties of glycerol enthalpies of combustion and vaporization and the heat capacity at 298.15 K. Enthalpies of solution in water at 288.15, 298.15, and 308.15 K", The Journal of Chemical Thermodynamics **20**, 11, 1353–1359 (1988).

Bates, R. G. and G. D. Pinching, "Acidic dissociation constant and related thermodynamic quantities for monoethanolammonium ion in water from 0 to 50°C", Journal of Research of the National Bureau of Standards **46**, 5, 349–352 (1951).

Belin, B. J., N. Busset, E. Giraud, A. Molinaro, A. Silipo and D. K. Newman, "Hopanoid lipids: from membranes to plant–bacteria interactions", Nature Reviews Microbiology **16**, 304–315 (2018).

Ben-Naim, A. and Y. Marcus, "Solvation thermodynamics of nonionic solutes", The Journal of Chemical Physics **81**, 4, 2016–2027 (1984).

Benning, C., Z.-H. Huang and D. A. Gage, "Accumulation of a novel glycolipid and a betaine lipid in cells of *Rhodobacter sphaeroides* grown under phosphate limitation", Archives of Biochemistry and Biophysics **317**, 1, 103–111 (1995).

Benson, S. W. and J. H. Buss, "Additivity rules for the estimation of molecular properties. Thermodynamic properties", The Journal of Chemical Physics **29**, 3, 546–572 (1958).

Berry, A., O. Harriott, R. Moreau, S. Osman, D. Benson and A. Jones, "Hopanoid lipids compose the *Frankia* vesicle envelope, presumptive barrier of oxygen diffusion to nitrogenase", Proceedings of the National Academy of Sciences **90**, 13, 6091–6094 (1993).

Blumenberg, M., M. Hoppert, M. Krüger, A. Dreier and V. Thiel, "Novel findings on hopanoid occurrences among sulfate reducing bacteria: Is there a direct link to nitrogen fixation?", Organic Geochemistry **49**, 1–5 (2012).

Boyd, E., T. Hamilton, J. Wang, L. He and C. Zhang, "The role of tetraether lipid composition in the adaptation of thermophilic archaea to acidity", Frontiers in Microbiology **4**, 62 (2013).

Boyd, E. S., A. Pearson, Y. Pi, W.-J. Li, Y. G. Zhang, L. He, C. L. Zhang and G. G. Geesey, "Temperature and pH controls on glycerol dibiphytanyl glycerol tetraether lipid composition in the hyperthermophilic crenarchaeon *Acidilobus sulfurireducens*", Extremophiles **15**, 1, 59–65 (2011).

Bradley, A. S., A. Pearson, J. P. Sáenz and C. J. Marx, "Adenosylhopane: the first intermediate in hopanoid side chain biosynthesis", Organic Geochemistry **41**, 10, 1075–1081 (2010).

Bringer, S., T. Härtner, K. Poralla and H. Sahm, "Influence of ethanol on the hopanoid content and the fatty acid pattern in batch and continuous cultures of *Zymomonas mobilis*", Archives of Microbiology **140**, 4, 312–316 (1985).

Brocks, J. J., R. Buick, G. A. Logan and R. E. Summons, "Composition and syngeneity of molecular fossils from the 2.78 to 2.45 billion-year-old Mount Bruce Supergroup, Pilbara Craton, Western Australia", Geochimica et Cosmochimica Acta **67**, 22, 4289–4319 (2003a).

Brocks, J. J., R. Buick, R. E. Summons and G. A. Logan, "A reconstruction of Archean biological diversity based on molecular fossils from the 2.78 to 2.45 billion-year-old Mount Bruce Supergroup, Hamersley Basin, Western Australia", Geochimica et Cosmochimica Acta **67**, 22, 4321–4335 (2003b).

Cabani, S., P. Gianni, V. Mollica and L. Lepori, "Group contributions to the thermodynamic properties of non-ionic organic solutes in dilute aqueous solution", Journal of Solution Chemistry **10**, 8, 563–595 (1981).

Cabani, S., V. Mollica, L. Lepori and S. Lobo, "Volume changes in the proton ionization of amines in water. 1. morpholines and piperazines", The Journal of Physical Chemistry **81**, 10, 982–987 (1977a).

Cabani, S., V. Mollica, L. Lepori and S. Lobo, "Volume changes in the proton ionization of amines in water. 2. Amino alcohols, amino ethers, and diamines", The Journal of Physical Chemistry **81**, 10, 987–993 (1977b).

Chao, J., K. R. Hall, K. N. Marsh and R. C. Wilhoit, "Thermodynamic properties of key organic oxygen compounds in the carbon range C1 to C4. Part 2. Ideal gas properties", Journal of Physical and Chemical Reference Data **15**, 4, 1369–1436 (1986).

Chavez, A. L. and G. G. Birch, "The hydrostatic and hydrodynamic volumes of polyols in aqueous solutions and their sweet taste", Chemical Senses **22**, 2, 149–161 (1997).

Cooke, M. P., H. M. Talbot and P. Farrimond, "Bacterial populations recorded in bacteriohopanepolyol distributions in soils from northern england", Organic Geochemistry **39**, 9, 1347–1358 (2008).

Cox, A., E. L. Shock and J. R. Havig, "The transition to microbial photosynthesis in hot spring ecosystems", Chemical Geology **280**, 3, 344–351 (2011).

Cox, J., D. D. Wagman and V. A. Medvedev, *CODATA key values for thermodynamics* (Chem/Mats-Sci/E, 1989).

da Silva, M. A. R., M. D. R. da Silva, A. I. L. Ferreira, Q. Shi, B. F. Woodfield and R. N. Goldberg, "Thermochemistry of α-d-xylose (cr)", The Journal of Chemical Thermodynamics **58**, 20–28 (2013).

Daniel, R. M. and D. A. Cowan, "Biomolecular stability and life at high temperatures", Cellular and Molecular Life Sciences **57**, 2, 250–264 (2000).

Deamer, D., J. P. Dworkin, S. A. Sandford, M. P. Bernstein and L. J. Allamandola, "The first cell membranes", Astrobiology **2**, 4, 371–381 (2002).

Dehn, W. M., "Comparative solubilities in water, in pyridine and in aqueous pyridine", Journal of the American Chemical Society **39**, 7, 1399–1404 (1917).

Di Lorenzo, F., I. Paciello, L. L. Fazio, L. Albuquerque, L. Sturiale, M. S. da Costa, R. Lanzetta, M. Parrilli, D. Garozzo, M. L. Bernardini, D. A. Silipo and A. Molinaro, "Thermophiles as potential source of novel endotoxin antagonists: the full structure and bioactivity of the lipo-oligosaccharide from *Thermomonas hydrothermalis*", ChemBioChem **15**, 14, 2146–2155 (2014).

Dick, J. M., "Calculation of the relative metastabilities of proteins using the CHNOSZ software package", Geochemical Transactions **9**, 1 (2008).

Dick, J. M., D. E. LaRowe and H. C. Helgeson, "Temperature, pressure, and electrochemical constraints on protein speciation: Group additivity calculation of the standard molal thermodynamic properties of ionized unfolded proteins", Biogeosciences **3**, 3, 311–336 (2006).

Dick, J. M. and E. L. Shock, "Calculation of the relative chemical stabilities of proteins as a function of temperature and redox chemistry in a hot spring", Public Library of Science One **6**, 8, e22782 (2011).

Dick, J. M. and E. L. Shock, "A metastable equilibrium model for the relative abundances of microbial phyla in a hot spring", Public Library of Science One **8**, 9, e72395 (2013).

Diercks, H., A. Semeniuk, N. Gisch, H. Moll, K. A. Duda and G. Hölzl, "Accumulation of novel glycolipids and ornithine lipids in *Mesorhizobium loti* under phosphate deprivation", Journal of Bacteriology **197**, 3, 497–509 (2015).

DiPaola, G. and B. Belleau, "Polyol–water interactions. Apparent molal heat capacities and volumes of aqueous polyol solutions", Canadian Journal of Chemistry **55**, 22, 3825–3830 (1977).

Domalski, E. S. and E. D. Hearing, "Estimation of the thermodynamic properties of C-H-N-O-S-halogen compounds at 298.15 K", Journal of Physical and Chemical Reference Data **22**, 4, 805–1159 (1993).

Ferreira, A. M., R. Wait, M. F. Nobre and M. S. da Costa, "Characterization of glycolipids from *Meiothermus* spp.", Microbiology **145**, 5, 1191–1199 (1999).

Grogan, D. W. and J. E. Cronan, "Cyclopropane ring formation in membrane lipids of bacteria", Microbiology and Molecular Biology Reviews **61**, 4, 429–441 (1997).

Guschina, I. A. and J. L. Harwood, "Lipids and lipid metabolism in eukaryotic algae", Progress in Lipid Research **45**, 2, 160–186 (2006).

Halverson, L. J. and M. K. Firestone, "Differential effects of permeating and nonpermeating solutes on the fatty acid composition of *Pseudomonas putida*", Applied and Environmental Microbiology **66**, 6, 2414–2421 (2000).

Hamborg, E. S. and G. F. Versteeg, "Dissociation constants and thermodynamic properties of amines and alkanolamines from (293 to 353) K", Journal of Chemical & Engineering Data **54**, 4, 1318–1328 (2009).

Havig, J. R., J. Raymond, D. R. Meyer-Dombard, N. Zolotova and E. L. Shock, "Merging isotopes and community genomics in a siliceous sinter-depositing hot spring", Journal of Geophysical Research: Biogeosciences **116**, G1 (2011).

Heipieper, H. J., G. Meulenbeld, Q. van Oirschot and J. De Bont, "Effect of environmental factors on the trans/cis ratio of unsaturated fatty acids in *Pseudomonas putida* S12", Applied and Environmental Microbiology **62**, 8, 2773–2777 (1996).

Jia, C., C. L. Zhang, W. Xie, J.-X. Wang, F. Li, S. Wang, H. Dong, W. Li and E. S. Boyd, "Differential temperature and pH controls on the abundance and composition of H-GDGTs in terrestrial hot springs", Organic Geochemistry **75**, 109–121 (2014).

Karlsson, A. Å., P. Michélsen and G. Odham, "Molecular species of sphingomyelin: Determination by high-performance liquid chromatography/mass spectrometry with electrospray and high-performance liquid chromatography/tandem mass spectrometry with atmospheric pressure chemical ionization", Journal of Mass Spectrometry **33**, 12, 1192–1198 (1998).

Kaur, G., B. W. Mountain, M. B. Stott, E. C. Hopmans and R. D. Pancost, "Temperature and pH control on lipid composition of silica sinters from diverse hot springs in the Taupo Volcanic Zone, New Zealand", Extremophiles **19**, 2, 327–344 (2015).

Khan, I. and P. Brimblecombe, "Henry's law constants of low molecular weight ($<$ 130) organic acids", Journal of Aerosol Science **23**, 897–900 (1992).

Kiran, M. D., S. Annapoorni, I. Suzuki, N. Murata and S. Shivaji, "Cis–trans isomerase gene in psychrophilic *Pseudomonas syringae* is constitutively expressed during growth and under conditions of temperature and solvent stress", Extremophiles **9**, 2, 117–125 (2005).

Koga, Y., T. Kyuragi, M. Nishihara and N. Sone, "Did archaeal and bacterial cells arise independently from noncellular precursors? a hypothesis stating that the advent of membrane phospholipid with enantiomeric glycerophosphate backbones

caused the separation of the two lines of descent", Journal of molecular evolution **46**, 1, 54–63 (1998).

Lian, Y., A. Chen, J. Suurkuusk and I. Wadso, "Polyol-water interactions as reflected by aqueous heat capacity values", Acta Chemica Scandinavica **36**, 9 (1982).

Lightfoot, D., A. Baron and J. Wootton, "Expression of the *Escherichia coli* glutamate dehydrogenase gene in the cyanobacterium *Synechococcus* PCC6301 causes ammonium tolerance", Plant Molecular Biology **11**, 3, 335–344 (1988).

Likens, G. E., *Biogeochemistry of inland waters* (Academic Press, 2010).

Liu, W., E. Sakr, P. Schaeffer, H. M. Talbot, J. Donisi, T. Härtner, E. Kannenberg, E. Takano and M. Rohmer, "Ribosylhopane, a novel bacterial hopanoid, as precursor of C35 bacteriohopanepolyols in *Streptomyces coelicolor* A3 (2)", ChemBioChem **15**, 14, 2156–2161 (2014).

Loiacono, S. T., D. R. Meyer-Dombard, J. R. Havig, A. T. Poret-Peterson, H. E. Hartnett and E. L. Shock, "Evidence for high-temperature in situ nifH transcription in an alkaline hot spring of Lower Geyser Basin, Yellowstone National Park", Environmental Microbiology **14**, 5, 1272–1283 (2012).

Lopes Jesus, A., L. I. Tomé, M. E. S. Eusébio and J. Redinha, "Determination of the enthalpy of solute- solvent interaction from the enthalpy of solution: Aqueous solutions of erythritol and l-threitol", The Journal of Physical Chemistry B **110**, 18, 9280–9285 (2006).

Luisi, P. L., *The Emergence of Life: From Chemical Origins to Synthetic Biology, Second Edition* (Cambridge University Press, 2016).

Maham, Y., T. T. Teng, L. G. Hepler and A. E. Mather, "Densities, excess molar volumes, and partial molar volumes for binary mixtures of water with monoethanolamine, diethanolamine, and triethanolamine from 25 to 80°C", Journal of Solution Chemistry **23**, 2, 195–205 (1994).

Marlowe, I., J. Green, A. Neal, S. Brassell, G. Eglinton and P. Course, "Long chain (n-C37–C39) alkenones in the *Prymnesiophyceae*. Distribution of alkenones and other lipids and their taxonomic significance", British Phycological Journal **19**, 3, 203–216 (1984).

Martin, P., B. A. Van Mooy, A. Heithoff and S. T. Dyhrman, "Phosphorus supply drives rapid turnover of membrane phospholipids in the diatom *Thalassiosira pseudonana*", The International Society for Microbial Ecology journal **5**, 6, 1057–1060 (2011).

Martin, W. and M. J. Russell, "On the origins of cells: a hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells", Philosophical Transactions of the Royal Society B: Biological Sciences **358**, 1429, 59–85 (2003).

McCollom, T. M. and W. Bach, "Thermodynamic constraints on hydrogen generation during serpentinization of ultramafic rocks", Geochimica et Cosmochimica Acta **73**, 3, 856–875 (2009).

Minnikin, D., H. Abdolrahimzadeh and J. Baddiley, "Replacement of acidic phospholipids by acidic glycolipids in *Pseudomonas diminuta*", Nature **249**, 5454, 268–269 (1974).

Moore, E. K., E. C. Hopmans, W. I. C. Rijpstra, L. Villanueva, S. N. Dedysh, I. S. Kulichevskaya, H. Wienk, F. Schoutsen and J. S. S. Damsté, "Novel mono-, di-, and trimethylornithine membrane lipids in northern wetland planctomycetes", Applied and Environmental Microbiology **79**, 22, 6874–6884 (2013).

Newman, D. K., C. Neubauer, J. N. Ricci, C.-H. Wu and A. Pearson, "Cellular and molecular biological approaches to interpreting ancient biomarkers", Annual Review of Earth and Planetary Sciences **44**, 493–522 (2016).

Nichols, N., R. Sköld, C. Spink and I. Wadsö, "Thermochemistry of solutions of biochemical model compounds 6. α, ω-dicarboxylic acids,-diamines, and-diols in aqueous solution", The Journal of Chemical Thermodynamics **8**, 10, 993–999 (1976).

Nishihara, M. and Y. Koga, "Extraction and composition of polar lipids from the archaebacterium, *Methanobacterium thermoautotrophicum*: effective extraction of tetraether lipids by an acidified solvent", The Journal of Biochemistry **101**, 4, 997–1005 (1987).

Ourisson, G. and P. Albrecht, "Hopanoids. 1. Geohopanoids: the most abundant natural products on earth?", Accounts of Chemical Research **25**, 9, 398–402 (1992).

Ourisson, G., P. Albrecht and M. Rohmer, "The hopanoids: palaeochemistry and biochemistry of a group of natural products", Pure and Applied Chemistry **51**, 4, 709–729 (1979).

Ourisson, G. and M. Rohmer, "Prokaryotic polyterpenes: Phylogenetic precursors of sterols", in "Current Topics in Membranes and Transport", edited by F. Bronner and A. Kleinteller, vol. 17, pp. 153–182 (Elsevier, 1982).

Pearson, A., Z. Huang, A. Ingalls, C. Romanek, J. Wiegel, K. Freeman, R. Smittenberg and C. Zhang, "Nonmarine crenarchaeol in Nevada hot springs", Applied and Environmental Microbiology **70**, 9, 5229–5237 (2004).

Pearson, A., Y. Pi, W. Zhao, W. Li, Y. Li, W. Inskeep, A. Perevalova, C. Romanek, S. Li and C. L. Zhang, "Factors controlling the distribution of archaeal tetraethers in terrestrial hot springs", Applied and Environmental Microbiology **74**, 11, 3523–3532 (2008).

Pliego, J. R. and J. M. Riveros, "New values for the absolute solvation free energy of univalent ions in aqueous solution", Chemical Physics Letters **332**, 5, 597–602 (2000).

Plyasunov, A., E. Shock and J. OConnell, "Corresponding-states correlations for estimating partial molar volumes of nonelectrolytes at infinite dilution in water over extended temperature and pressure ranges", Fluid Phase Equilibria **247**, 1, 18–31 (2006).

Plyasunov, A. V., N. V. Plyasunova and E. L. Shock, "Group contribution values for the thermodynamic functions of hydration of aliphatic esters at 298.15 K, 0.1 MPa", Journal of Chemical & Engineering Data **49**, 5, 1152–1167 (2004).

Plyasunov, A. V. and E. L. Shock, "Thermodynamic functions of hydration of hydrocarbons at 298.15 K and 0.1 MPa", Geochimica et Cosmochimica Acta **64**, 3, 439–468 (2000).

Plyasunov, A. V. and E. L. Shock, "Correlation strategy for determining the parameters of the revised Helgeson-Kirkham-Flowers model for aqueous nonelectrolytes", Geochimica et Cosmochimica Acta **65**, 21, 3879–3900 (2001).

Plyasunova, N., A. Plyasunov and E. Shock, "Database of thermodynamic properties for aqueous organic compounds", International Journal of Thermophysics **25**, 2, 351–360 (2004).

Poger, D. and A. E. Mark, "The relative effect of sterols and hopanoids on lipid bilayers: when comparable is not identical", The Journal of Physical Chemistry B **117**, 50, 16129–16140 (2013).

Poralla, K., T. Härtner and E. Kannenberg, "Effect of temperature and ph on the hopanoid content of *Bacillus acidocaldarius*", FEMS Microbiology Letters **23**, 2-3, 253–256 (1984).

Poralla, K., E. Kannenberg and A. Blume, "A glycolipid containing hopane isolated from the acidophilic, thermophilic *Bacillus acidocaldarius*, has a cholesterol-like function in membranes", Federation of European Biochemical Societies letters **113**, 1, 107–110 (1980).

Poralla, K., G. Muth and T. Härtner, "Hopanoids are formed during transition from substrate to aerial hyphae in *Streptomyces coelicolor* A3 (2)", FEMS Microbiology Letters **189**, 1, 93–95 (2000).

Rashby, S. E., A. L. Sessions, R. E. Summons and D. K. Newman, "Biosynthesis of 2-methylbacteriohopanepolyols by an anoxygenic phototroph", Proceedings of the National Academy of Sciences **104**, 38, 15099–15104 (2007).

Ricci, J. N., M. L. Coleman, P. V. Welander, A. L. Sessions, R. E. Summons, J. R. Spear and D. K. Newman, "Diverse capacity for 2-methylhopanoid production correlates with a specific ecological niche", The ISME journal **8**, 3, 675–684 (2014).

Riekhof, W. R., S. Naik, H. Bertrand, C. Benning and D. R. Voelker, "Phosphate starvation in fungi induces the replacement of phosphatidylcholine with the phosphorus-free betaine lipid diacylglyceryl-n, n, n-trimethylhomoserine", Eukaryotic Cell **13**, 6, 749–757 (2014).

Rohmer, M., P. Bouvier-Nave and G. Ourisson, "Distribution of hopanoid triterpenes in prokaryotes", Microbiology **130**, 5, 1137–1150 (1984).

Sato, N., "Roles of the acidic lipids sulfoquinovosyl diacylglycerol and phosphatidylglycerol in photosynthesis: their specificity and evolution", Journal of Plant Research **117**, 6, 495–505 (2004).

Schaeffer, P., G. Schmitt, P. Adam and M. Rohmer, "Acid-catalyzed formation of 32, 35-anhydrobacteriohopanetetrol from bacteriohopanetetrol", Organic Geochemistry **39**, 10, 1479–1482 (2008).

Schmerk, C. L., M. A. Bernards and M. A. Valvano, "Hopanoid production is required for low-pH tolerance, antimicrobial resistance, and motility in *Burkholderia cenocepacia*", Journal of Bacteriology **193**, 23, 6712–6723 (2011).

Schouten, S., E. C. Hopmans and J. S. S. Damsté, "The organic geochemistry of glycerol dialkyl glycerol tetraether lipids: a review", Organic geochemistry **54**, 19–61 (2013).

Schouten, S., E. C. Hopmans, R. D. Pancost and J. S. S. Damsté, "Widespread occurrence of structurally diverse tetraether membrane lipids: evidence for the ubiquitous presence of low-temperature relatives of hyperthermophiles", Proceedings of the National Academy of Sciences **97**, 26, 14421–14426 (2000).

Schouten, S., E. C. Hopmans, E. Schefuß and J. S. S. Damste, "Distributional variations in marine crenarchaeotal membrane lipids: a new tool for reconstructing ancient sea water temperatures?", Earth and Planetary Science Letters **204**, 1-2, 265–274 (2002).

Schröder, J. M., *Intact polar lipids in marine sediments: improving analytical protocols and assessing planktonic and benthic sources*, Ph.D. thesis, Staats-und Universitätsbibliothek Bremen (2015).

Schubotz, F., L. E. Hays, R. D'Arcy, A. Gillespie, E. L. Shock and R. E. Summons, "Stable isotope labeling confirms mixotrophic nature of streamer biofilm communities at alkaline hot springs", Frontiers in Microbiology **6** (2015).

Schubotz, F., D. Meyer-Dombard, A. Bradley, H. Fredricks, K. Hinrichs, E. Shock and R. Summons, "Spatial and temporal variability of biomarkers and microbial diversity reveal metabolic and community flexibility in Streamer Biofilm Communities in the Lower Geyser Basin, Yellowstone National Park", Geobiology **11**, 6, 549–569 (2013).

Schubotz, F., S. G. Wakeham, J. S. Lipp, H. F. Fredricks and K.-U. Hinrichs, "Detection of microbial biomass by intact polar membrane lipid analysis in the water column and surface sediments of the Black Sea", Environmental Microbiology **11**, 10, 2720–2734 (2009).

Sedlbauer, J., "Database on the infinite dilution partial molar Gibbs energies of hydration of aqueous organic nonelectrolytes", URL:http://www.kch.tul.cz/texty/sedlbauer/gorg.nb (Updated 2012).

Sedlbauer, J., G. Bergin and V. Majer, "Group contribution method for Henry's Law constant of aqueous hydrocarbons", American Institute of Chemical Engineers Journal **48**, 12, 2936–2959 (2002).

Shock, E. and P. Canovas, "The potential for abiotic organic synthesis and biosynthesis at seafloor hydrothermal systems", Geofluids **10**, 1-2, 161–192 (2010).

Shock, E. L., "Organic acids in hydrothermal solutions: standard molal thermodynamic properties of carboxylic acids and estimates of dissociation constants at high temperatures and pressures", American Journal of Science **295**, 5, 496–580 (1995).

Shock, E. L. and H. C. Helgeson, "Calculation of the thermodynamic and transport properties of aqueous species at high pressures and temperatures: Correlation algorithms for ionic species and equation of state predictions to 5 kb and 1000°C", Geochimica et Cosmochimica Acta **52**, 8, 2009–2036 (1988).

Shock, E. L. and H. C. Helgeson, "Calculation of the thermodynamic and transport properties of aqueous species at high pressures and temperatures: Standard partial molal properties of organic species", Geochimica et Cosmochimica Acta **54**, 4, 915–945 (1990).

Shock, E. L., E. H. Oelkers, J. W. Johnson, D. A. Sverjensky and H. C. Helgeson, "Calculation of the thermodynamic properties of aqueous species at high pressures and temperatures. effective electrostatic radii, dissociation constants and standard partial molal properties to 1000°C and 5 kbar", Journal of the Chemical Society, Faraday Transactions **88**, 6, 803–826 (1992).

Silipo, A., C. De Castro, R. Lanzetta, M. Parrilli and A. Molinaro, "Lipopolysaccharides", in "Prokaryotic Cell Wall Compounds: Structure and Biochemistry", edited by H. König, H. Claus and A. Varma, pp. 133–153 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2010).

Sojo, V., A. Pomiankowski and N. Lane, "A bioenergetic basis for membrane divergence in archaea and bacteria", Public Library of Science Biology **12**, 8, e1001926 (2014).

Sollich, M., M. Y. Yoshinaga, S. Häusler, R. E. Price, K.-U. Hinrichs and S. I. Bühring, "Heat stress dictates microbial lipid composition along a thermal gradient in marine sediments", Frontiers in Microbiology **8**, 1550 (2017).

Sturt, H. F., R. E. Summons, K. Smith, M. Elvert and K.-U. Hinrichs, "Intact polar membrane lipids in prokaryotes and sediments deciphered by high-performance liquid chromatography/electrospray ionization multistage mass spectrometry-new biomarkers for biogeochemistry and microbial ecology", Rapid Communications in Mass Spectrometry **18**, 6, 617–628 (2004).

Sugden, S., "CXLII.–A relation between surface tension, density, and chemical composition", Journal of the Chemical Society, Transactions **125**, 1177–1189 (1924).

Summons, R. E., L. L. Jahnke, J. M. Hope and G. A. Logan, "2-Methylhopanoids as biomarkers for cyanobacterial oxygenic photosynthesis", Nature **400**, 6744, 554–557 (1999).

Sverjensky, D. A., B. Harrison and D. Azzolini, "Water in the deep Earth: the dielectric constant and the solubilities of quartz and corundum to 60 kb and 1200 °C", Geochimica et Cosmochimica Acta **129**, 125–145 (2014).

Talbot, H. M., P. Farrimond, P. Schaeffer and R. D. Pancost, "Bacteriohopanepolyols in hydrothermal vent biogenic silicates", Organic Geochemistry **36**, 4, 663–672 (2005).

Talbot, H. M., M. Rohmer and P. Farrimond, "Rapid structural elucidation of composite bacterial hopanoids by atmospheric pressure chemical ionisation liquid chromatography/ion trap mass spectrometry", Rapid Communications in Mass Spectrometry **21**, 6, 880–892 (2007a).

Talbot, H. M., M. Rohmer and P. Farrimond, "Structural characterisation of unsaturated bacterial hopanoids by atmospheric pressure chemical ionisation liquid chromatography/ion trap mass spectrometry", Rapid Communications in Mass Spectrometry **21**, 10, 1613–1622 (2007b).

Talbot, H. M., A. H. Squier, B. J. Keely and P. Farrimond, "Atmospheric pressure chemical ionisation reversed-phase liquid chromatography/ion trap mass spectrometry of intact bacteriohopanepolyols", Rapid Communications in Mass Spectrometry **17**, 7, 728–737 (2003a).

Talbot, H. M., R. Summons, L. Jahnke and P. Farrimond, "Characteristic fragmentation of bacteriohopanepolyols during atmospheric pressure chemical ionisation liquid chromatography/ion trap mass spectrometry", Rapid Communications in Mass Spectrometry **17**, 24, 2788–2796 (2003b).

Talbot, H. M., R. E. Summons, L. L. Jahnke, C. S. Cockell, M. Rohmer and P. Farrimond, "Cyanobacterial bacteriohopanepolyol signatures from cultures and natural environmental settings", Organic Geochemistry **39**, 2, 232–263 (2008).

Tanger, J. C. and H. C. Helgeson, "Calculation of the thermodynamic and transport properties of aqueous species at high pressures and temperatures; revised equations of state for the standard partial molal properties of ions and electrolytes", American Journal of Science **288**, 1, 19–98 (1988).

Tewari, Y. B. and R. N. Goldberg, "Thermodynamics of reactions catalyzed byliditol 2-dehydrogenase: the xylose assimilation pathway", The Journal of Chemical Thermodynamics **28**, 10, 1127–1144 (1996).

Tierney, J. E., "GDGT thermometry: Lipid tools for reconstructing paleotemperatures", The Paleontological Society Papers **18**, 115–132 (2012).

van Meer, G., D. R. Voelker and G. W. Feigenson, "Membrane lipids: where they are and how they behave", Nature Reviews Molecular Cell Biology **9**, 2, 112 (2008).

van Mooy, B. A., H. F. Fredricks, B. E. Pedler, S. T. Dyhrman, D. M. Karl, M. Koblížek, M. W. Lomas, T. J. Mincer, L. R. Moore, T. Moutin *et al.*, "Phytoplankton in the ocean use non-phosphorus lipids in response to phosphorus scarcity", Nature **458**, 7234, 69–72 (2009).

van Winden, J. F., H. M. Talbot, N. Kip, G.-J. Reichart, A. Pol, N. P. McNamara, M. S. Jetten, H. J. O. den Camp and J. S. S. Damste, "Bacteriohopanepolyol signatures as markers for methanotrophic bacteria in peat moss", Geochimica et Cosmochimica Acta **77**, 52–61 (2012).

Verevkin, S. P., "Determination of vapor pressures and enthalpies of vaporization of 1, 2-alkanediols", Fluid Phase Equilibria **224**, 1, 23–29 (2004).

Verevkin, S. P., V. N. Emelyanenko and G. Nell, "1, 2-propanediol. comprehensive experimental and theoretical study", The Journal of Chemical Thermodynamics **41**, 10, 1125–1131 (2009).

Verevkin, S. P., D. H. Zaitsau, V. N. Emelyanenko and A. A. Zhabina, "Thermodynamic properties of glycerol: Experimental and theoretical study", Fluid Phase Equilibria **397**, 87–94 (2015).

Wada, H. and N. Murata, "Lipids in thylakoid membranes and photosynthetic cells", in "Lipids in Photosynthesis", edited by H. Wada and N. Murata, pp. 1–9 (Springer, 2009).

Wang, M., G. H. Kim, F. Wei, H. Chen, J. Altarejos and X. Han, "Improved method for quantitative analysis of methylated phosphatidylethanolamine species and its application for analysis of diabetic-mouse liver samples", Analytical and Bioanalytical Chemistry **407**, 17, 5021–5032 (2015).

Wang, Z., Q. Wang, X. Liu, W. Fang, Y. Li and H. Xiao, "Measurement and correlation of solubility of xylitol in binary water + ethanol solvent mixtures between 278.00 K and 323.00 K", Korean Journal of Chemical Engineering **30**, 4, 931–936 (2013).

Weijers, J. W., S. Schouten, E. C. Hopmans, J. A. Geenevasen, O. R. David, J. M. Coleman, R. D. Pancost and J. S. Sinninghe Damsté, "Membrane lipids of mesophilic anaerobic bacteria thriving in peats have typical archaeal traits", Environmental Microbiology **8**, 4, 648–657 (2006).

Welander, P. V., D. M. Doughty, C.-H. WU, S. Mehay, R. E. Summons and D. K. Newman, "Identification and characterization of *Rhodopseudomonas palustris* TIE-1 hopanoid biosynthesis mutants", Geobiology **10**, 2, 163–177 (2012).

Welander, P. V., R. C. Hunter, L. Zhang, A. L. Sessions, R. E. Summons and D. K. Newman, "Hopanoids play a role in membrane integrity and pH homeostasis in *Rhodopseudomonas palustris* TIE-1", Journal of Bacteriology **191**, 19, 6145–6156 (2009).

White, D. C. and D. B. Ringelberg, "Signature lipid biomarker analysis", in "Techniques in microbial ecology", edited by R. S. Burlage, R. Atlas, D. Stahl, G. Geesey and G. Sayler, vol. 255, pp. 255–272 (Oxford University Press: New York, 1998).

Wolery, T., "Eq3/6-software for geochemical modeling, version 8.0", Tech. rep., UCRL-CODE-2003-009, Lawrence Livermore National Laboratory, Livermore CA (2002).

Wörmer, L., J. S. Lipp, J. M. Schröder and K.-U. Hinrichs, "Application of two new LC–ESI–MS methods for improved detection of intact polar lipids (IPLs) in environmental samples", Organic Geochemistry **59**, 10–21 (2013).

Wu, W., C. Zhang, H. Wang, L. He, W. Li and H. Dong, "Impacts of temperature and pH on the distribution of archaeal lipids in Yunnan hot springs, China", Frontiers in Microbiology **4**, 312 (2013).

Yang, Y.-L., F.-L. Yang, S.-C. Jao, M.-Y. Chen, S.-S. Tsay, W. Zou and S.-H. Wu, "Structural elucidation of phosphoglycolipids from strains of the bacterial thermophiles *Thermus* and *Meiothermus*", Journal of Lipid Research **47**, 8, 1823–1832 (2006).

Yaws, C. L., *Handbook of chemical compound data for process safety* (Gulf Professional Publishing, 1997).

Yoshinaga, M. Y., M. Y. Kellermann, P. E. Rossel, F. Schubotz, J. S. Lipp and K.-U. Hinrichs, "Systematic fragmentation patterns of archaeal intact polar lipids by high-performance liquid chromatography/electrospray ionization ion-trap mass spectrometry", Rapid Communications in Mass Spectrometry **25**, 23, 3563–3574 (2011).

Zhang, X., S. M. Ferguson-Miller and G. E. Reid, "Characterization of ornithine and glutamine lipids extracted from cell membranes of *Rhodobacter sphaeroides*", Journal of the American Society for Mass Spectrometry **20**, 2, 198–212 (2009).

Zhang, Y.-M. and C. O. Rock, "Membrane lipid homeostasis in bacteria", Nature Reviews Microbiology **6**, 3, 222–233 (2008).

Zhu, K., D. O. Bayles, A. Xiong, R. Jayaswal and B. J. Wilkinson, "Precursor and temperature modulation of fatty acid composition and growth of *Listeria monocytogenes* cold-sensitive mutants with transposon-interrupted branched-chain α-keto acid dehydrogenase", Microbiology **151**, 2, 615–623 (2005).

APPENDIX A

INTACT POLAR LIPID RESPONSE FACTORS AND MASS SPECTRAL INTERPRETATION OF NOVEL STRUCTURES

Analytical response factors. Table A.1 provides information about response factors for the suite of intact polar lipid (IPL) standards used in this work. Response factors were interpreted as the linear slopes of high-performance liquid chromatography-mass spectrometry (HPLC-MS) integrated peak area as a function of injected mass with a y-intercept of zero. The squared correlation coefficient, R^2, is reported for each regression. Two sets of response factors are given: one for a sample batch run in 2013 and another in 2014. Samples MS1, MS3, MS4, and MS5 use the response factors for 2014, while all other samples use those for 2013.

Mass spectral interpretation of novel structures. Mass spectra and putative chemical structures of lipids identified in this work are shown in Figures A.1 through A.9. All lipid structures shown in this appendix are tentatively assigned based on MS/MS fragmentation patterns. Carbon positions of alkyl chain modifications, bonding, and configuration between glycolipid headgroup moieties, *etc.*, are guesses. Dotted arrows refer to inferred fragmentation of the displayed structure and point in the direction of the fragment with the monoisotopic mass, in Da, listed at the end of the arrow. pNLC stands for 'precursor neutral loss chromatogram'. This a chromatogram based on the difference between the mass of a lipid precursor ion before and after the loss of a diagnostic fragment. Usually, in the case of IPLs, this is the headgroup or a piece of the headgroup. M+H and M-H refer to the monoisotopic mass of the displayed structure plus or minus a hydrogen atom, respectively, and assumes the resulting fragment has an overall charge of +1.

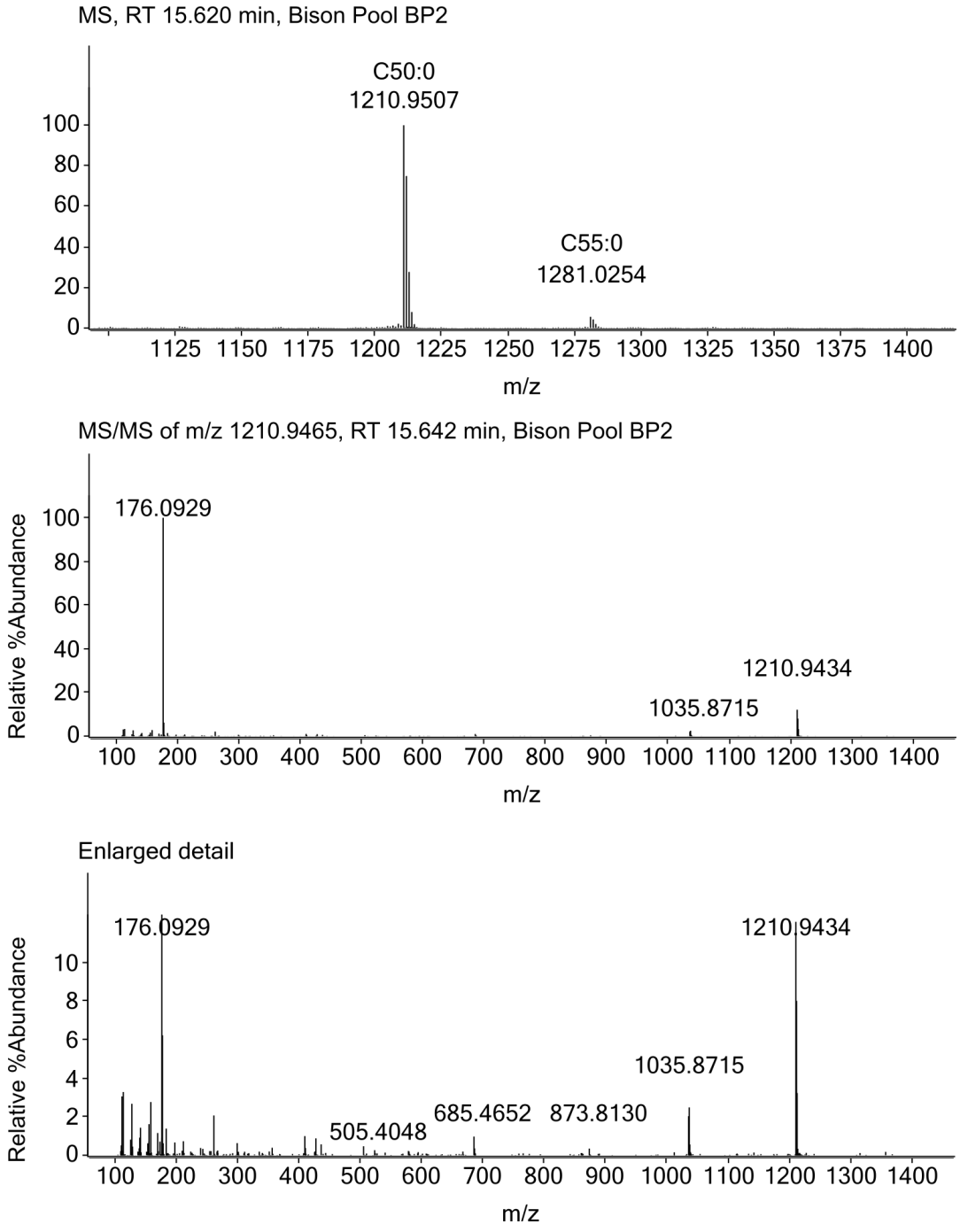
Table A.1: HPLC-MS IPL standards and response factors

			Batch 1[b]		Batch 2[c]	
No.[a]	Standard	Chains	RF	R^2	RF	R^2
1	1Gly-DAG	mix of C34:6, C36:6	6.37×10^5	0.95	9.45×10^5	0.99
2	1Gly-GDGT-PG	mix of 0-3 rings, ∼20% H-shaped	6.72×10^4	0.96	1.16×10^4	0.91
3	PC-DAG	C42:0	7.31×10^5	0.99	2.71×10^5	0.91
4	2Gly-DAG	mix of C34:2, C34:3 C34:6, C36:6	1.33×10^5	0.99	9.90×10^4	0.99
5	SQ-DAG	mix of C34:2, C34:3, C36:6	1.31×10^5	0.95	1.90×10^5	0.99
6	PI-DAG	C32:0	7.94×10^3	0.88	5.95×10^3	0.99
7	DPG	C72:4	3.97×10^5	0.95	7.24×10^4	0.99
8	PDME-DAG	C32:0	1.11×10^6	0.97	1.96×10^6	0.99
9	PE-DAG	C32:0	4.06×10^5	0.99	2.92×10^5	0.95
10	PG-DAG	C32:0	1.49×10^5	0.98	1.02×10^5	0.99
11	PME-DAG	C32:0	5.45×10^5	0.94	5.93×10^5	0.99
12	PS	C32:0	2.03×10^4	0.98	2.30×10^4	0.97
13	DGTS-d9	C32:0	1.65×10^6	0.89	7.16×10^6	0.96

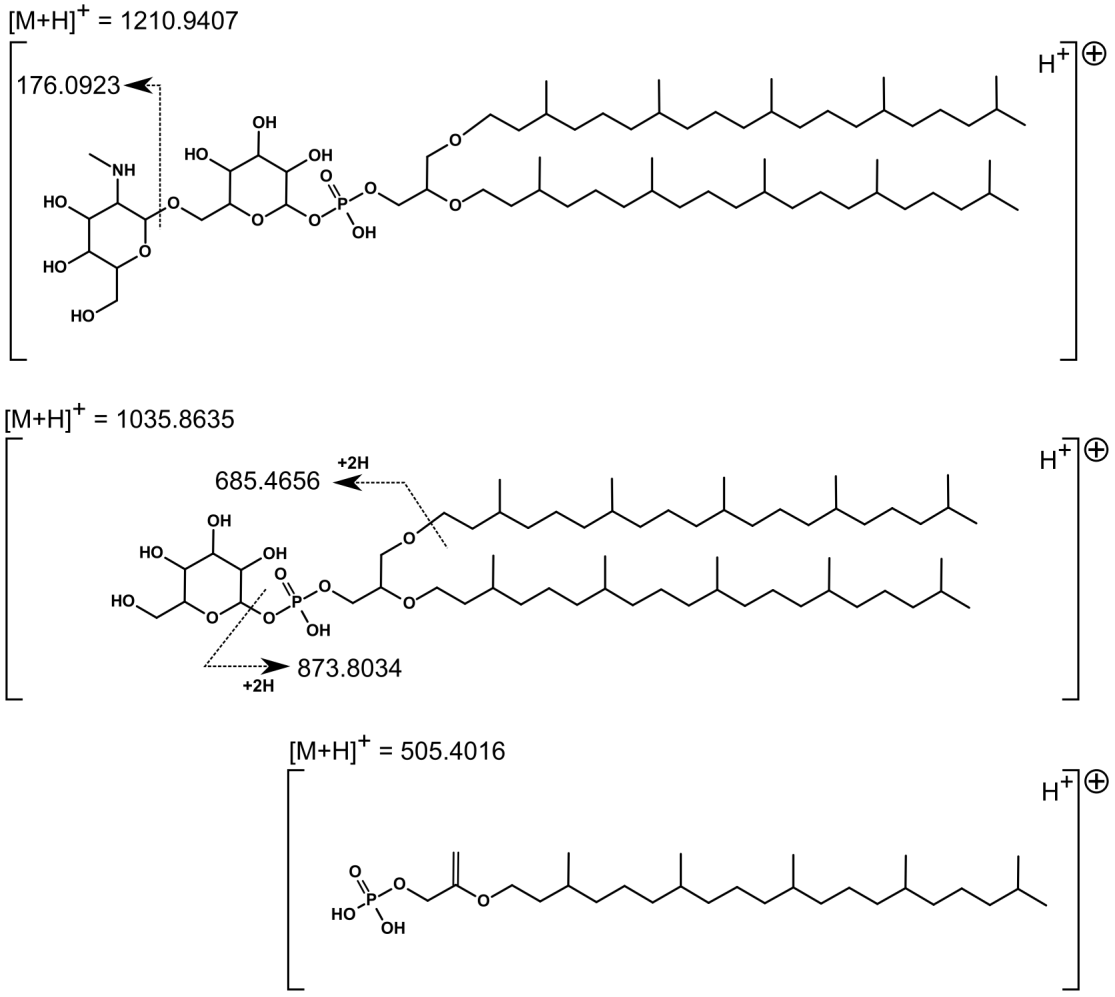
[a] Corresponds to numbered standards in Table 2.1 of main text.

[b] Batch 1 was analyzed in 2013 and includes all Bison Pool, Empress Pool, and Octopus Spring samples, as well as Mound Spring sample MS2. Linear response factors and R^2 were taken from two replicate calibrations of 0.1, 0.5, 1, 5, 10, and 50 ng standard injections.

[c] Batch 2 was analyzed in 2014 and includes Mound Spring samples MS1, MS3, MS4, and MS5. Linear response factors and R^2 were taken from a single calibration of 0.1, 0.5, 1, 5, and 10 ng standard injections.



(a) mass spectra



(b) putative fragments

Figure A.1: Mass spectra and putative structure for (N-methyl)glycosaminyl monoglycosyl phosphatidylarchaeol (MeNG-G-P-AR).

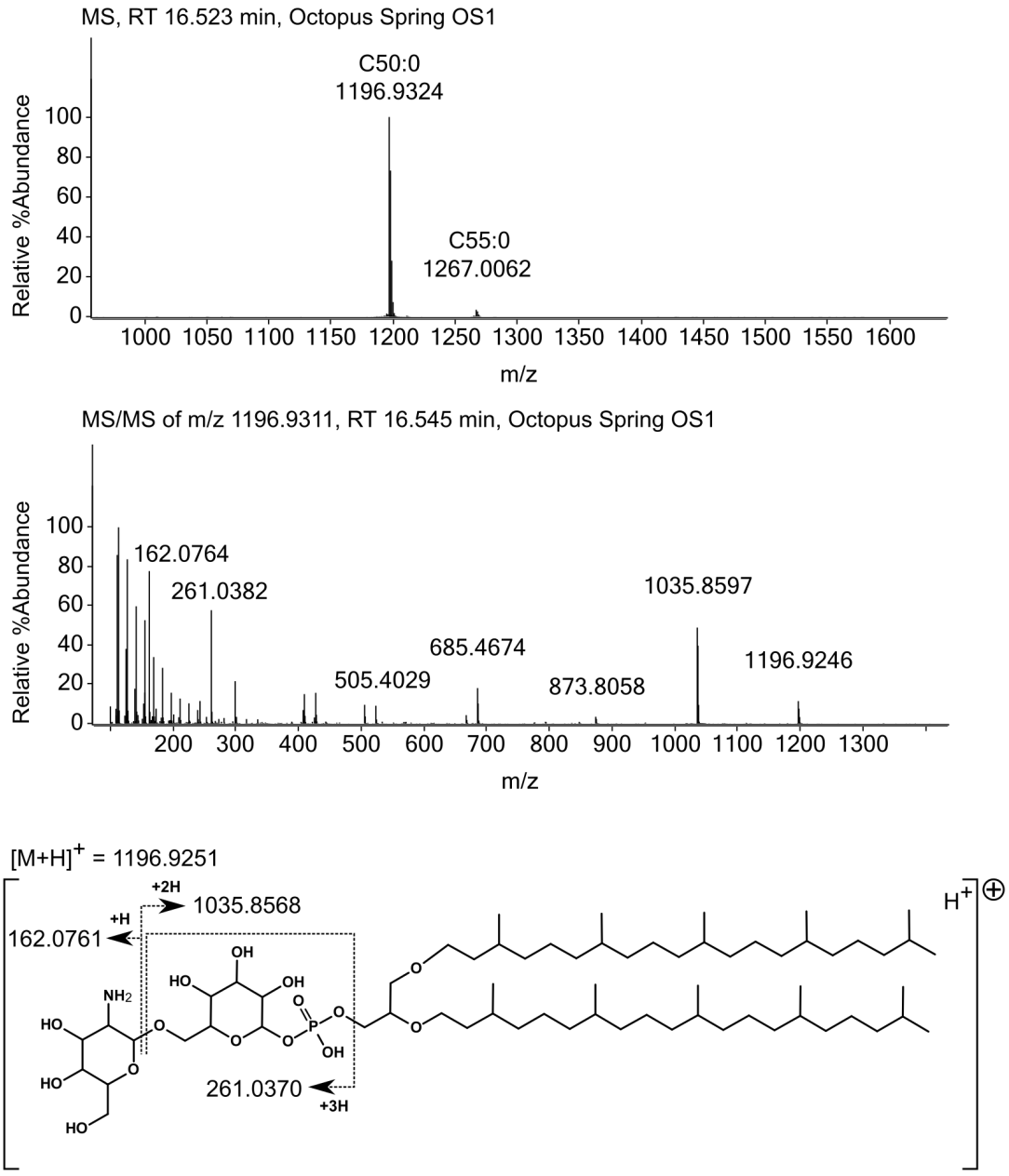


Figure A.2: Mass spectra and putative structure for (N)glycosaminyl monoglycosyl phosphatidylarchaeol (NG-G-P-AR). This and MeNG-G-P-AR share many fragments in MS/MS due to their structural similarity (see Figure A.1).

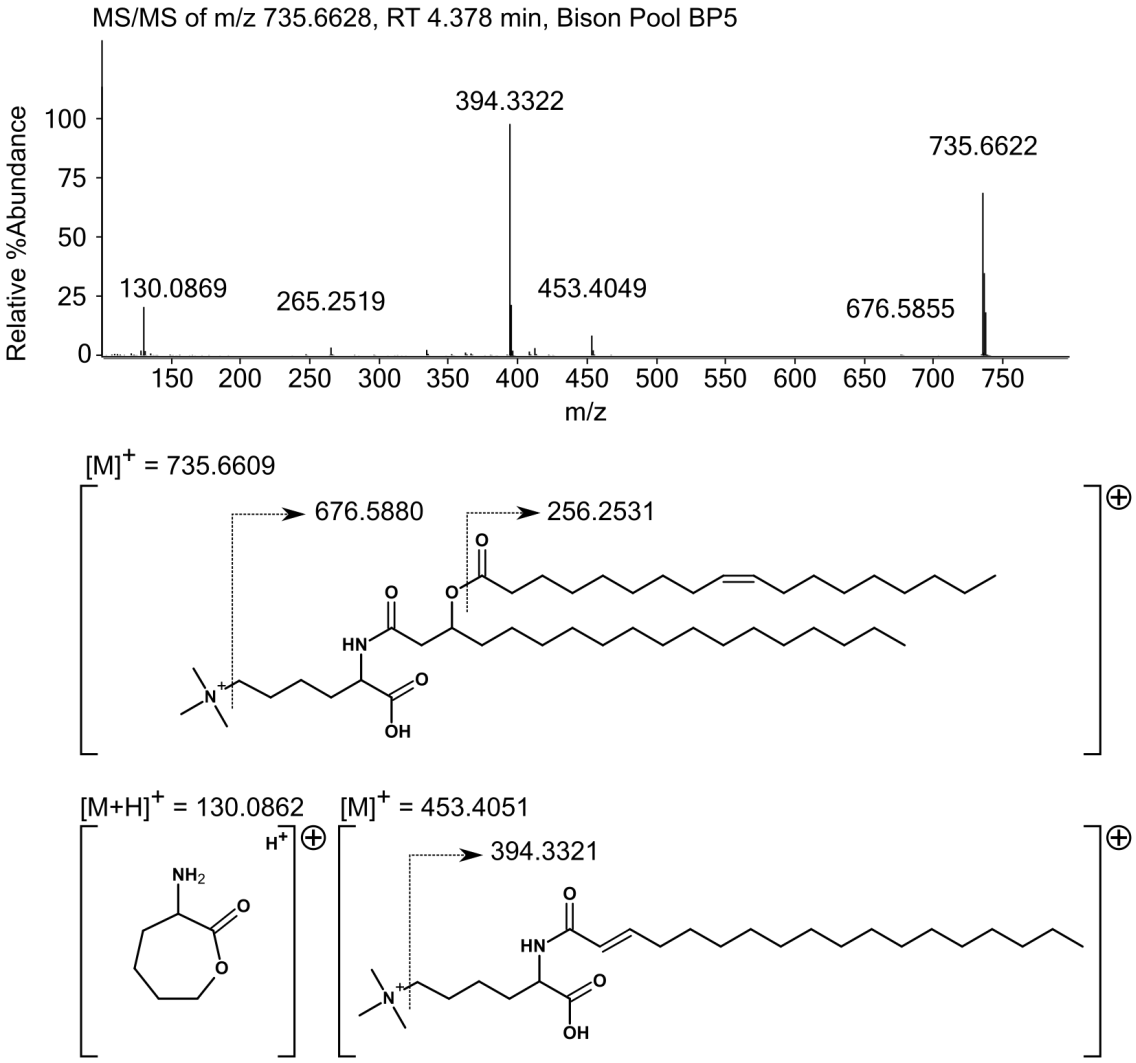


Figure A.3: Mass spectrum and putative structure for (6-N,6-N,6-N)trimethyllysine lipid (TM-KL).

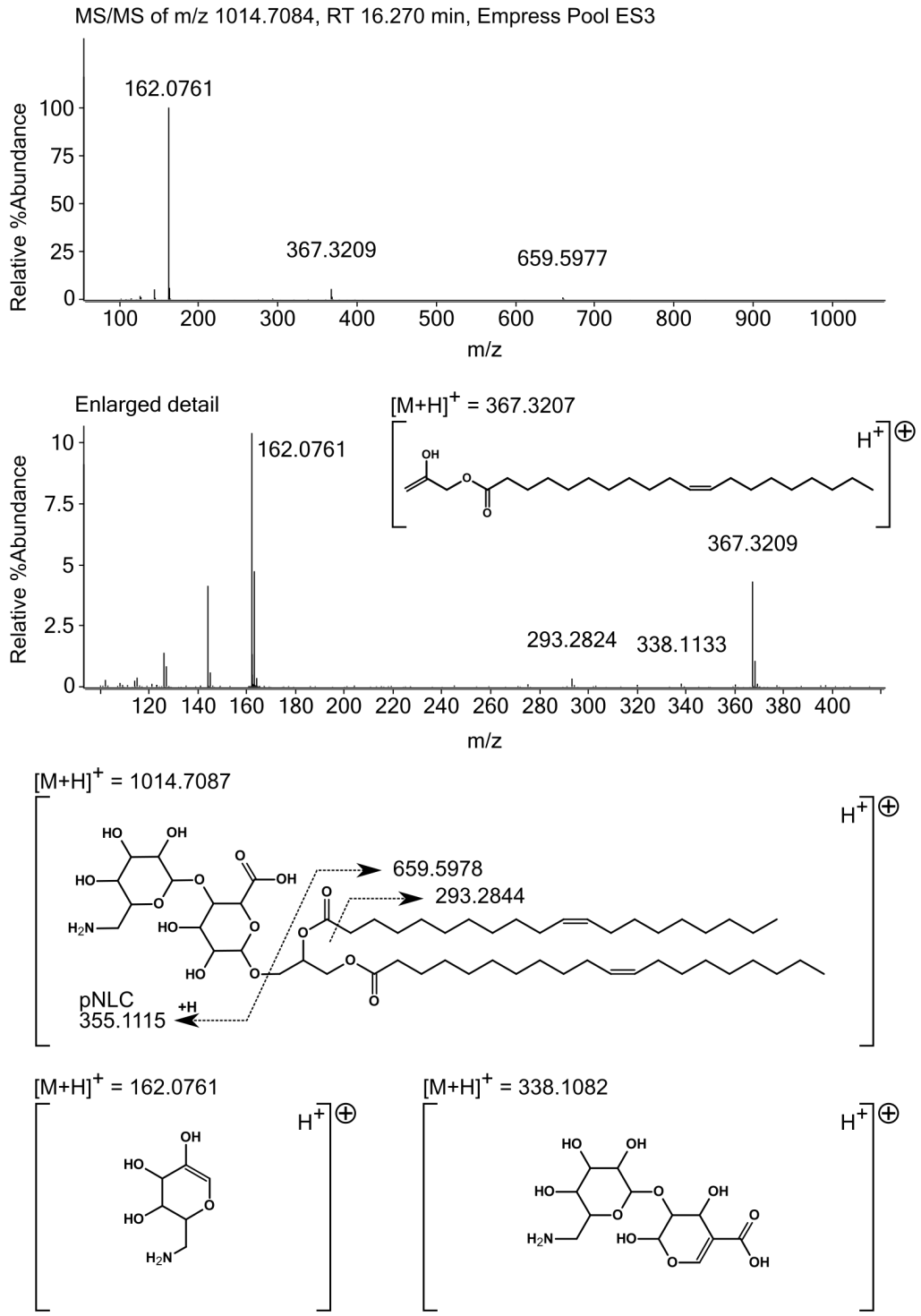
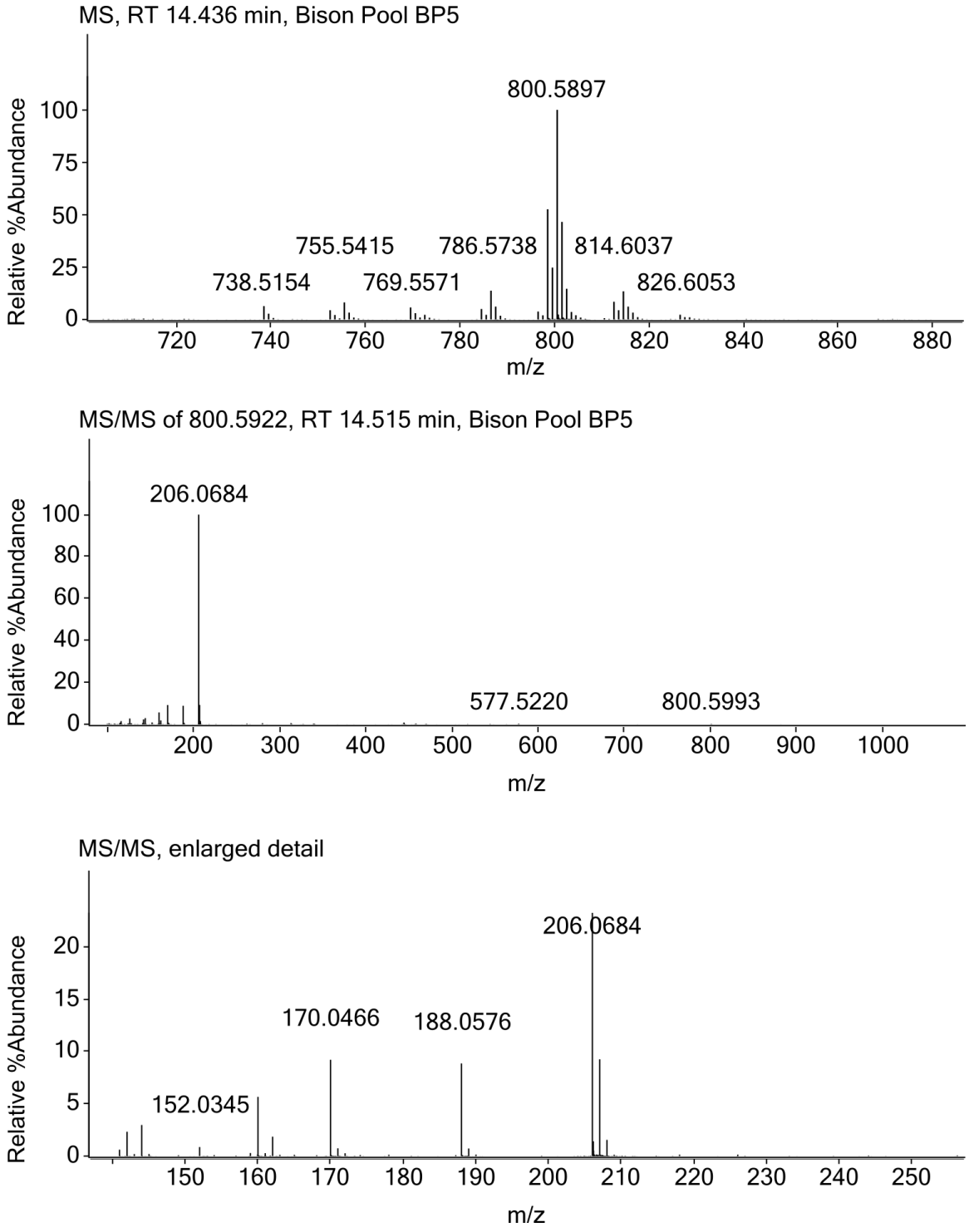
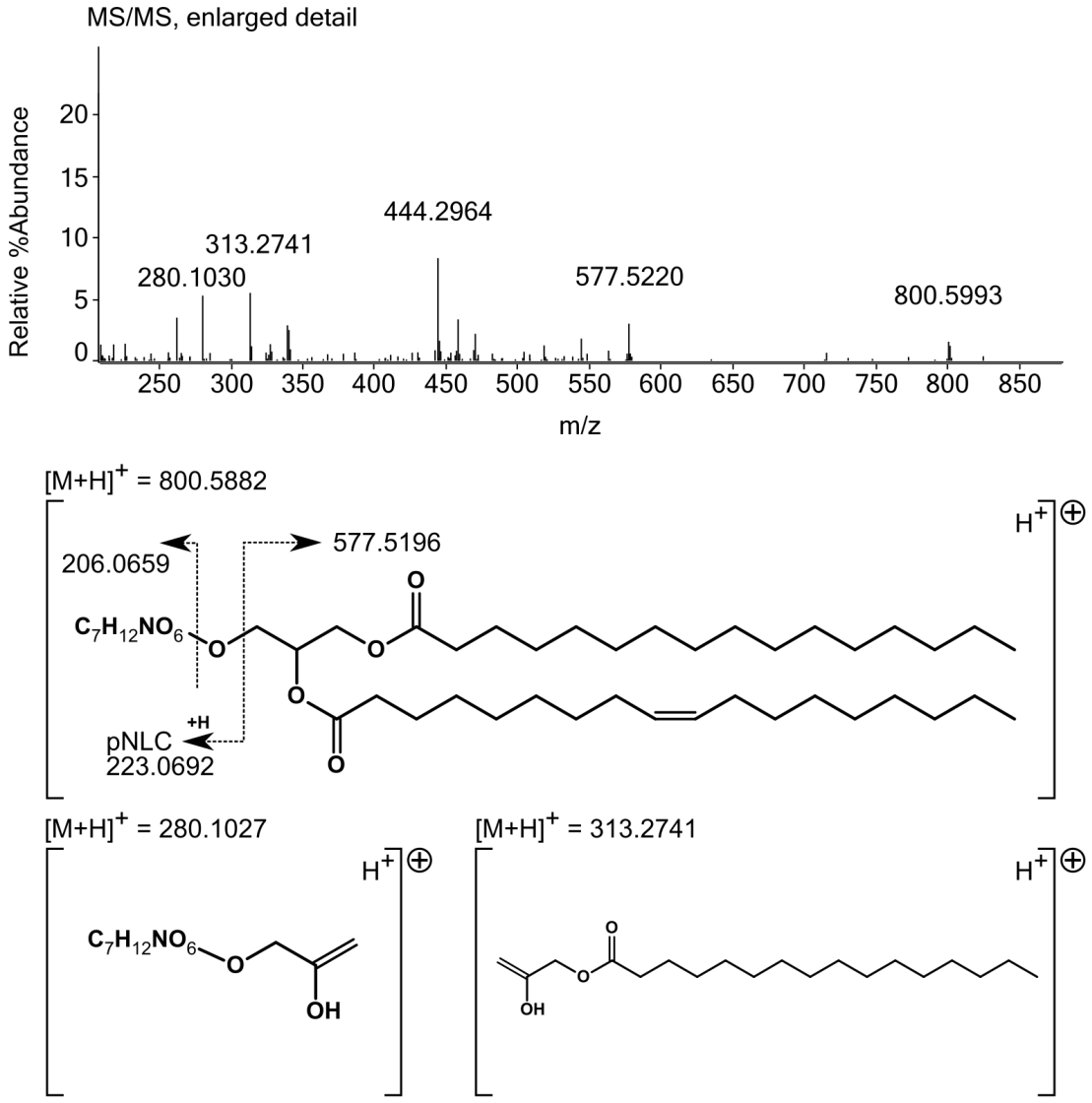


Figure A.4: Mass spectra and putative structure for (6-N,6-N,6-N)trimethyllysine lipid (NG-GA-DAG).



(a) The MS/MS spectrum of C34:1 223-DAG within the m/z 250 800 range. Neutral loss of the headgroup results in [M + H]+ - m/z 223.07 = m/z 577.51, corresponding to the ion shown below, where nC in acyl chains R1 and R2 sum to 34 with one unsaturation.



(b) The MS/MS spectrum of C34:1 223-DAG within the m/z 140 250 range. The entire height of the 206 headgroup fragment is not shown. Dehydration reactions of base peak 206 results in: m/z 206.07 18.01 = m/z 188.06; m/z 206.07 2*18.01 = m/z 170.05; m/z 206.07 3*18.01 = m/z 152.04.

Figure A.5: Mass spectra and putative structure for unknown lipid '223'-DAG.

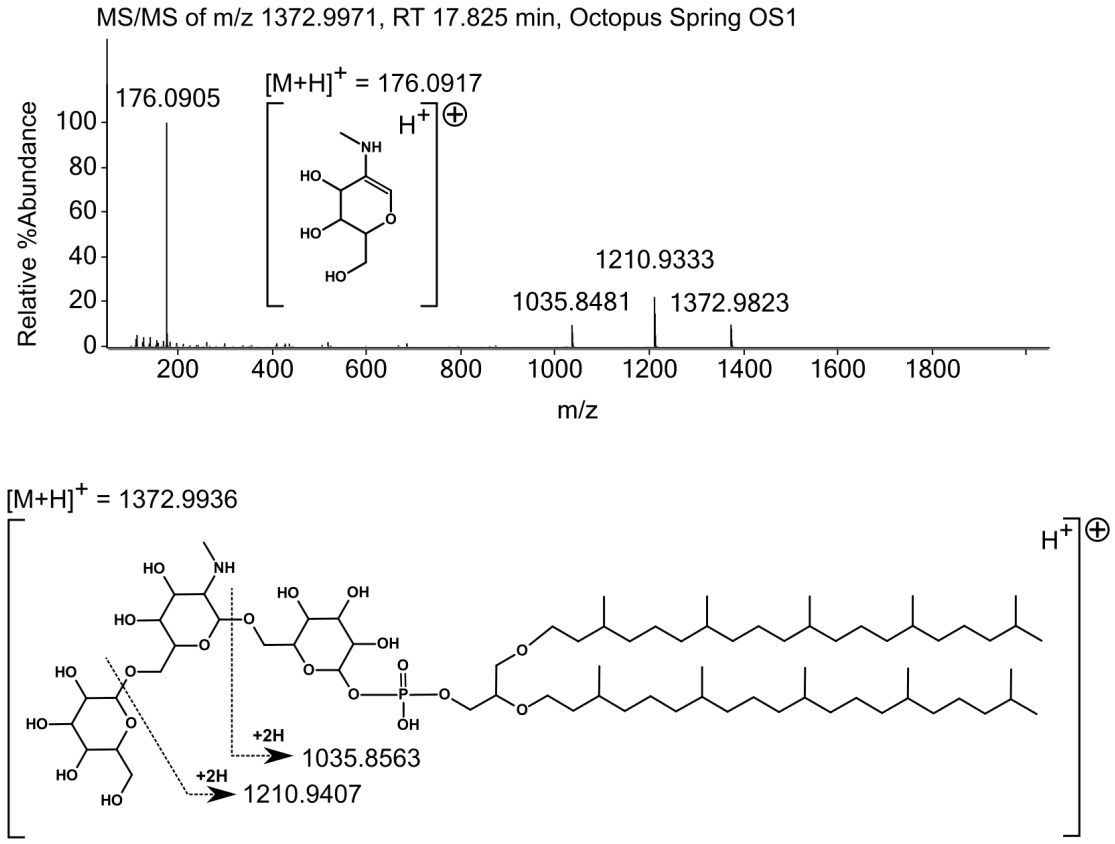


Figure A.6: Mass spectrum and putative structure for glycosyl (N-methyl)glycosaminyl glycosyl phosphatidylarchaeol (G-MeNG-G-P-AR).

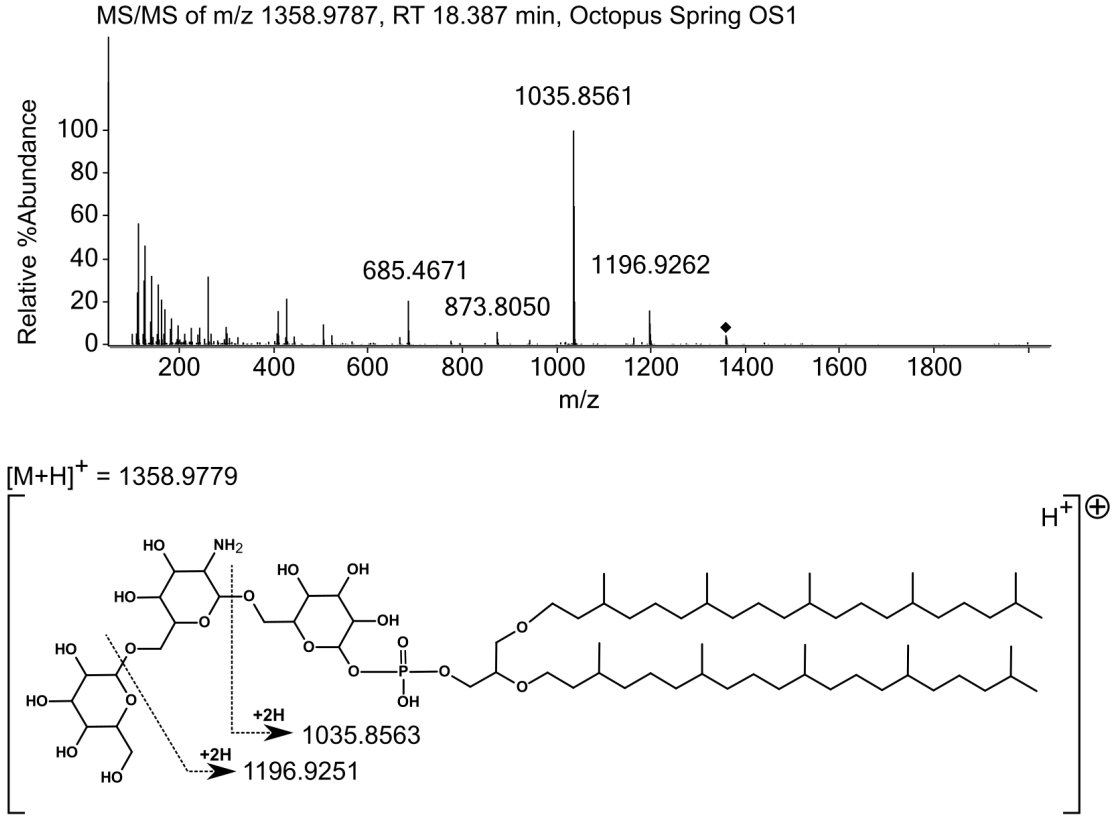
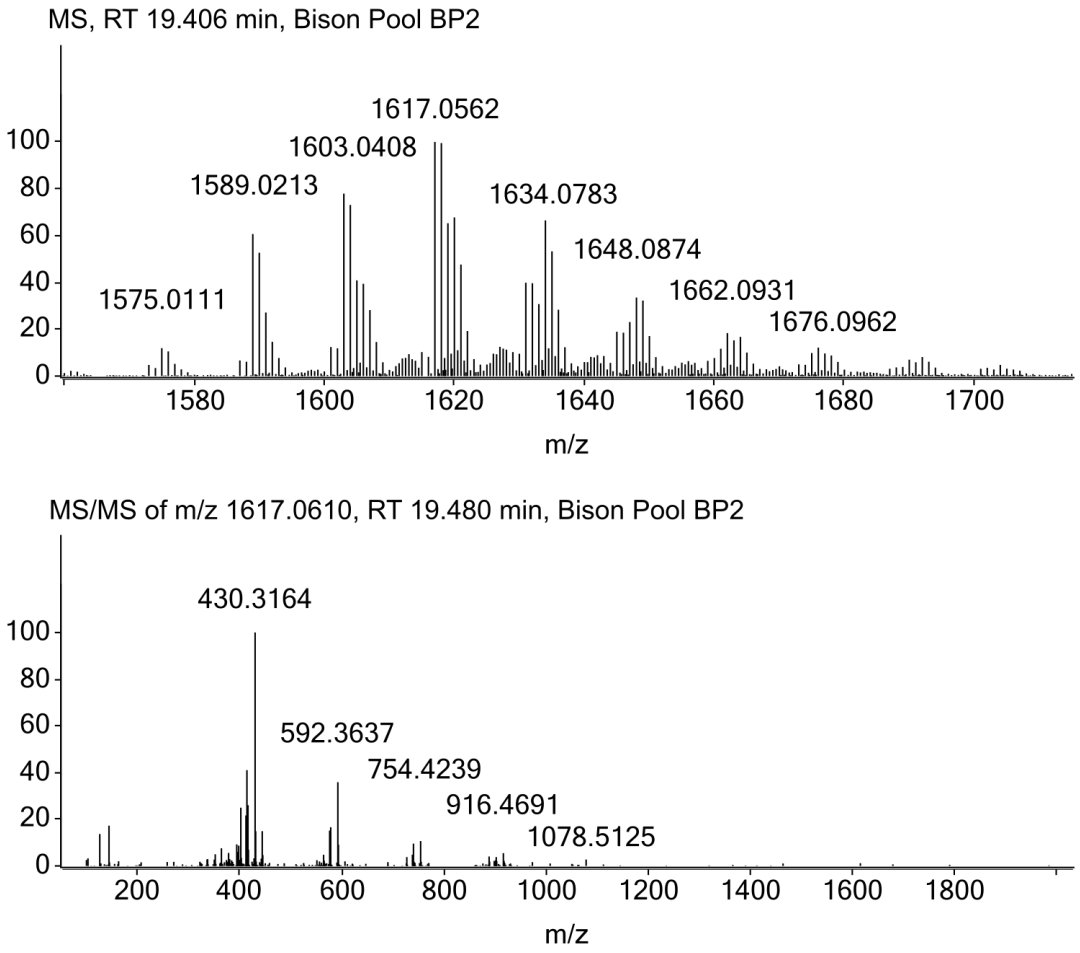
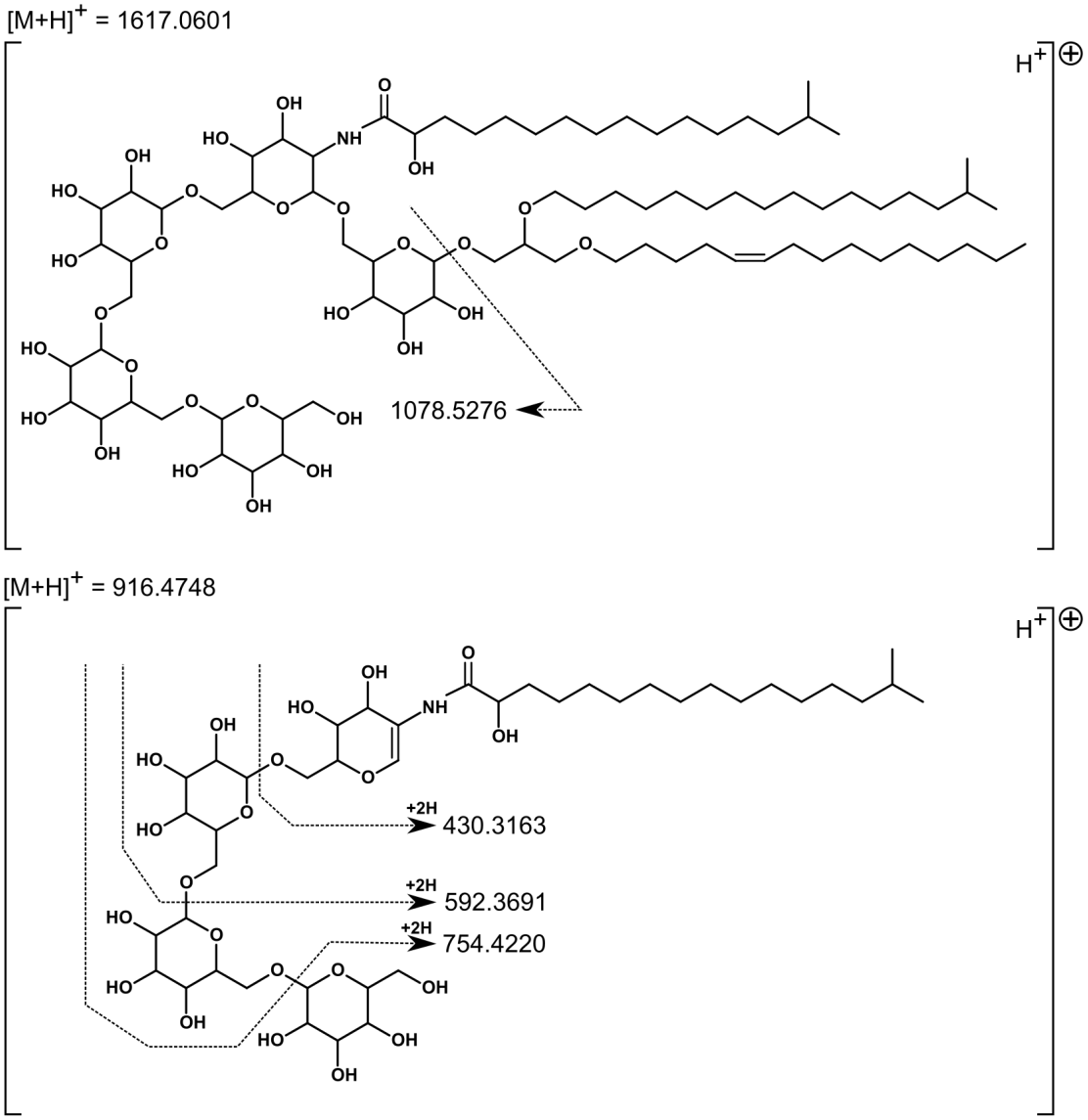


Figure A.7: Mass spectra and putative structure for glycosyl (N)glycosaminyl glycosyl phosphatidylarchaeol (G-MeNG-G-P-AR). See Figure A.1 for putative structures for fragments m/z 685 and 873.


MS, RT 19.406 min, Bison Pool BP2
1617.0562
1603.0408
1589.0213
1634.0783
1648.0874
1662.0931
1676.0962
1575.0111
100
80
60
40
20
0
1580
1600
1620
1640
1660
1680
1700
m/z
MS/MS of m/z 1617.0610, RT 19.480 min, Bison Pool BP2
430.3164
592.3637
754.4239
916.4691
1078.5125
100
80
60
40
20
0
200
400
600
800
1000
1200
1400
1600
1800
m/z

(a)



(b)

Figure A.8: (a) Mass spectra and (b) putative structure for triglycosyl (N-acetyl)glycosaminyl glycosyl dietherglycerol (3G-NAcG-G-DEG).

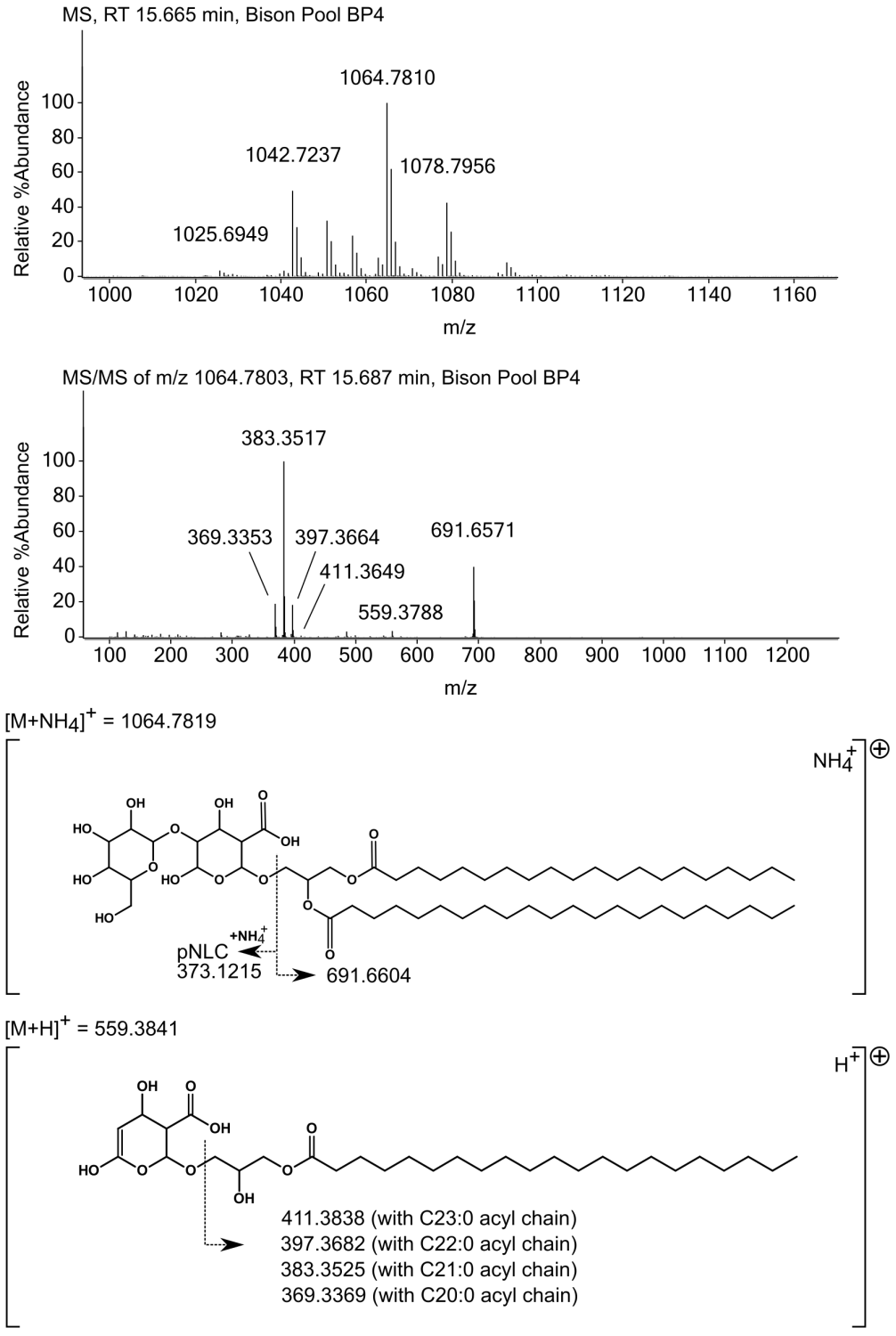


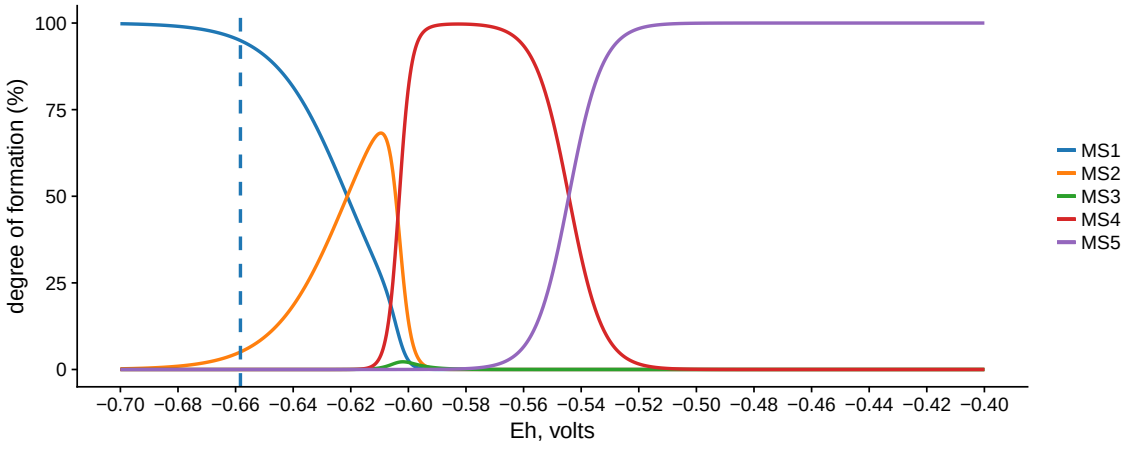
Figure A.9: Mass spectra and putative structure for G-GA-DAG.

APPENDIX B

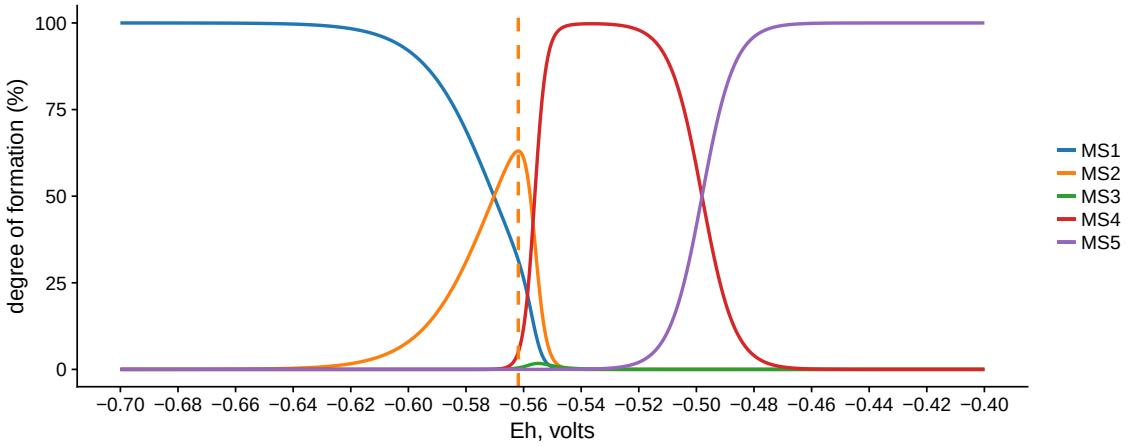
SUPPLEMENTARY FIGURES: PREDICTED EQUILIBRIUM ABUNDANCES AND SENSITIVITY ANALYSES OF SITE-AVERAGE FREE ALKYL CHAINS

Equilibrium abundances. Predicted equilibrium abundances of aqueous sample-averaged free alkyl chains as a function of Eh are shown in Figure B.1 for Mound Spring, Figure B.2 for Empress Pool, and Figure B.3 for Octopus Spring. Each subfigure shows a calculation at the temperature and basis species concentrations of a sample collected along the outflow channel of a spring. Dotted vertical lines indicate the alkyl chain-predicted Eh, Eh_{alkyl}, of the sample. Figures for Bison Pool samples can be found in Chapter 3.

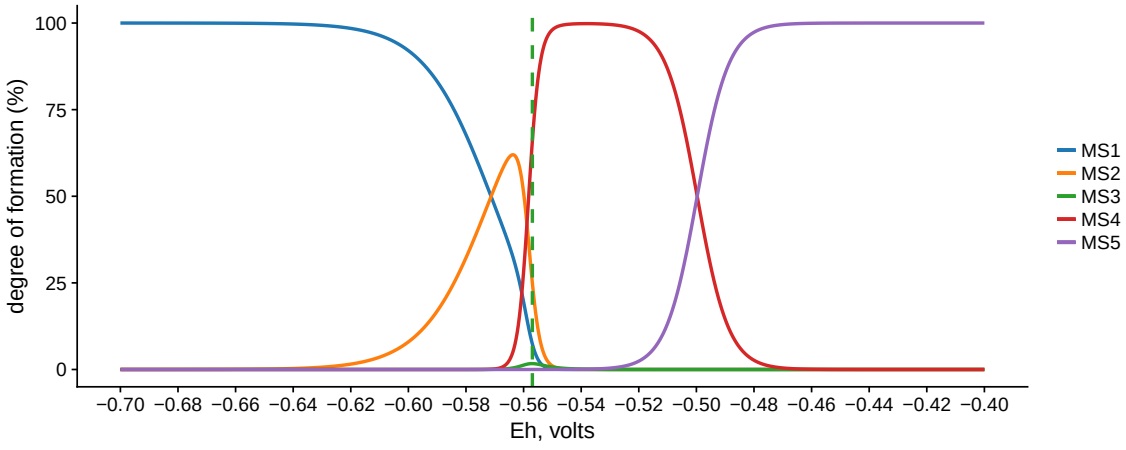
Sensitivity analysis. Distributions in predicted equilibrium abundance of aqueous sample-averaged free alkyl chains as a function of Eh produced by the sensitivity analysis are shown in Figure B.4 for Mound Spring, Figure B.5 for Empress Pool, and Figure B.6 for Octopus Spring. Sensitivity analysis involved 999 iterations of Monte Carlo-style random variation of lipid HPLC peak areas by 30% and mass spectral response factors between 0.01x and 100x. Results have a resolution of 400 calculations per iteration, binned here in 0.02 volt increments. Distributions for each sample-averaged free alkyl chain are shown as a colored box-and-whisker within each bin. Dark horizontal lines indicate median values for each distribution, with 50% of the results around the median falling within the colored box. Whiskers extend to an observation 1.5 times the IQR beyond this range. Values that fall outside the span of the whiskers are not indicated to reduce visual clutter. Figures for Bison Pool samples can be found in Chapter 3.


100
75
50
25
0
degree of formation (%)
−0.70 −0.68 −0.66 −0.64 −0.62 −0.60 −0.58 −0.56 −0.54 −0.52 −0.50 −0.48 −0.46 −0.44 −0.42 −0.40
Eh, volts
MS1
MS2
MS3
MS4
MS5

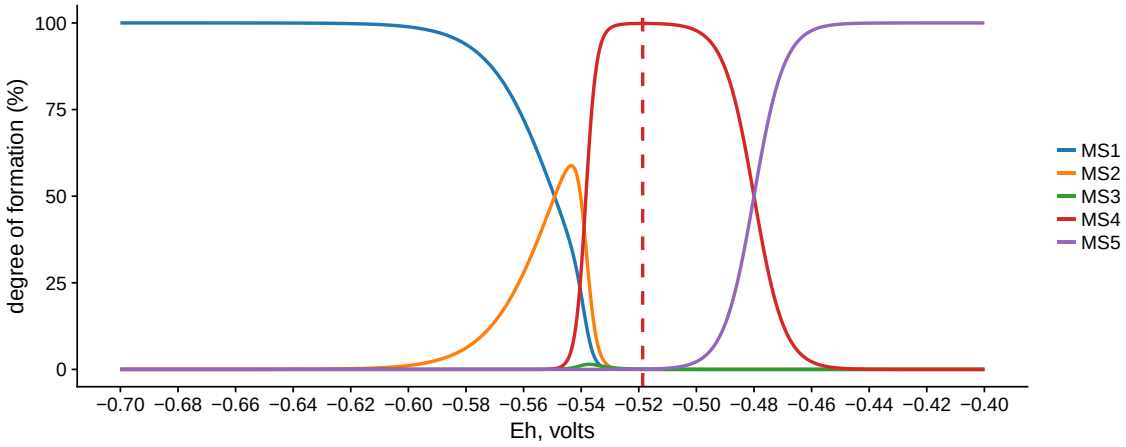
(a) MS1


100
75
50
25
0
degree of formation (%)
−0.70 −0.68 −0.66 −0.64 −0.62 −0.60 −0.58 −0.56 −0.54 −0.52 −0.50 −0.48 −0.46 −0.44 −0.42 −0.40
Eh, volts
MS1
MS2
MS3
MS4
MS5

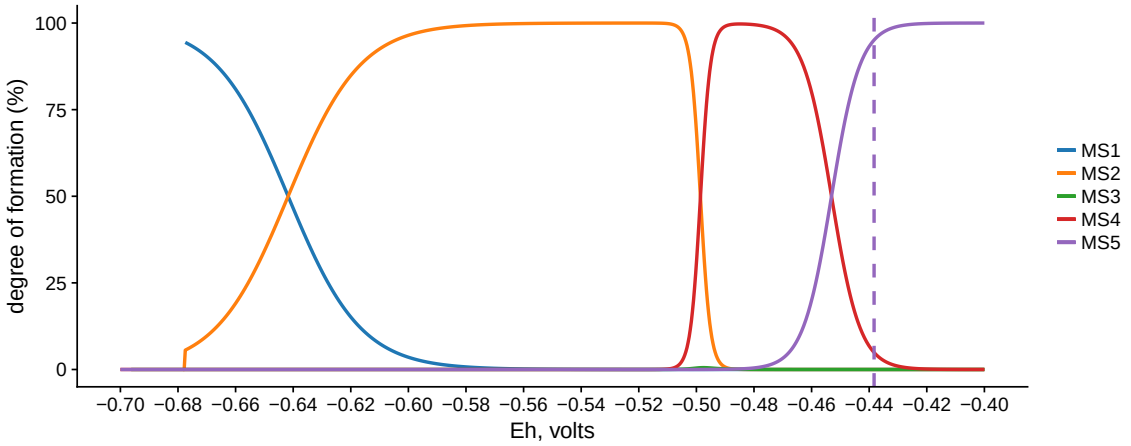
(b) MS2


100
75
50
25
0
degree of formation (%)
−0.70 −0.68 −0.66 −0.64 −0.62 −0.60 −0.58 −0.56 −0.54 −0.52 −0.50 −0.48 −0.46 −0.44 −0.42 −0.40
Eh, volts
MS1
MS2
MS3
MS4
MS5

(c) MS3

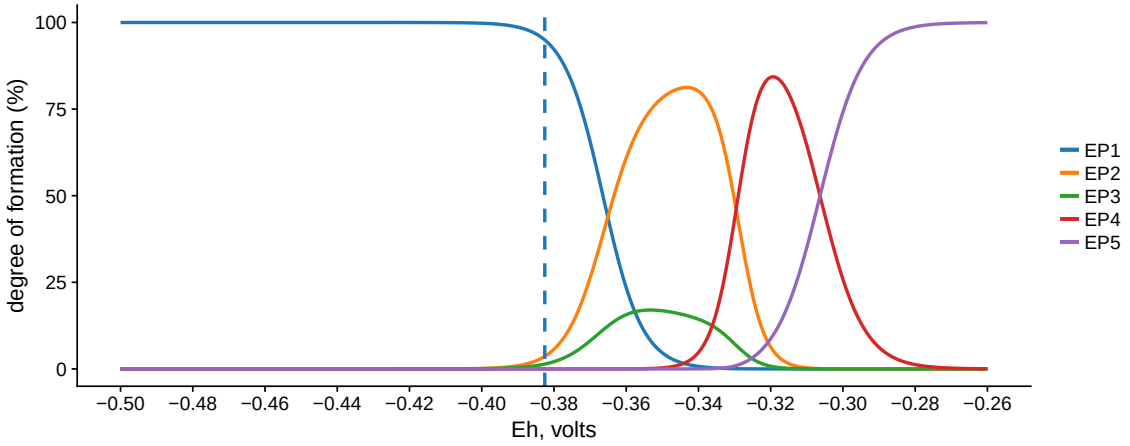

100
75
50
25
0
degree of formation (%)
−0.70 −0.68 −0.66 −0.64 −0.62 −0.60 −0.58 −0.56 −0.54 −0.52 −0.50 −0.48 −0.46 −0.44 −0.42 −0.40
Eh, volts
MS1
MS2
MS3
MS4
MS5

(d) MS4

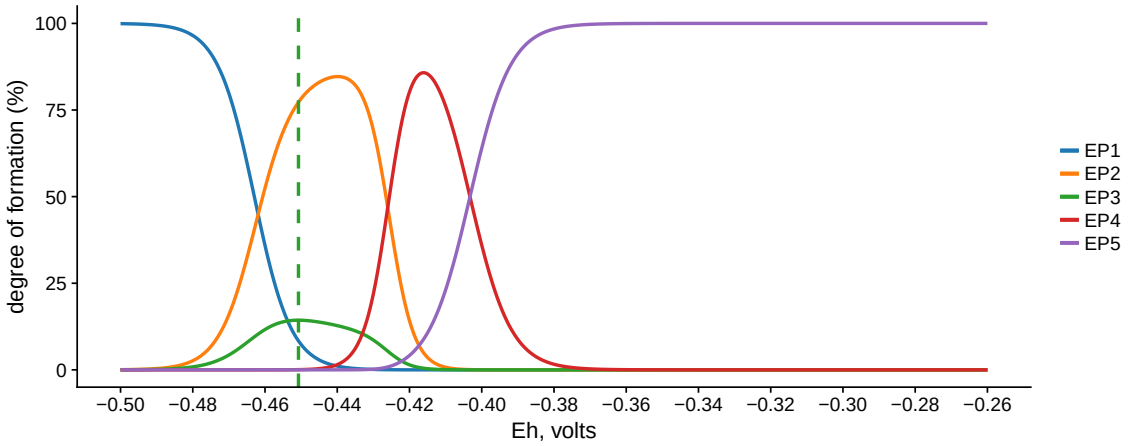


(e) MS5

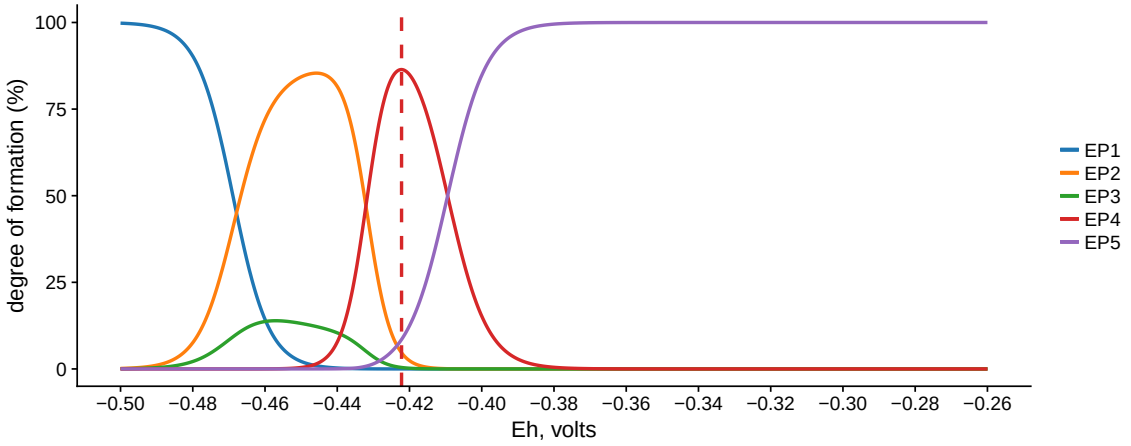
Figure B.1: Metastable equilibrium abundance of sample-averaged free alkyl chains of Mound Spring predicted across an Eh gradient at the temperatures and bicarbonate, ammonium (or nitrate for MS5), and proton concentrations measured at (a) MS1, (b) MS2, (c) MS3, (d) MS4, and (e) MS5. The vertical dotted line indicates the alkyl chain-predicted $\mathrm{Eh}_{\mathrm{alkyl}}$ of the sample.


degree of formation (%)
100
75
50
25
0
−0.50 −0.48 −0.46 −0.44 −0.42 −0.40 −0.38 −0.36 −0.34 −0.32 −0.30 −0.28 −0.26
Eh, volts
EP1
EP2
EP3
EP4
EP5

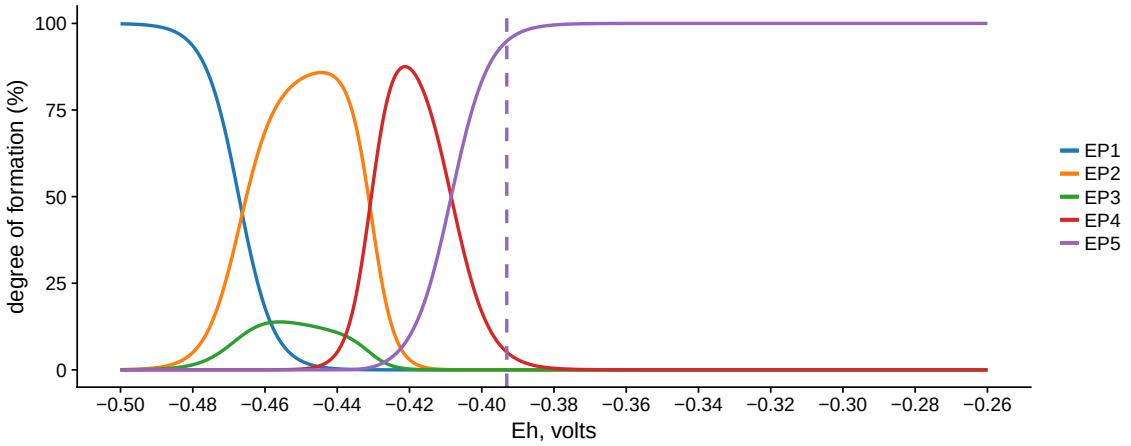
(a) EP1


degree of formation (%)
100
75
50
25
0
−0.50 −0.48 −0.46 −0.44 −0.42 −0.40 −0.38 −0.36 −0.34 −0.32 −0.30 −0.28 −0.26
Eh, volts
EP1
EP2
EP3
EP4
EP5

(b) EP3

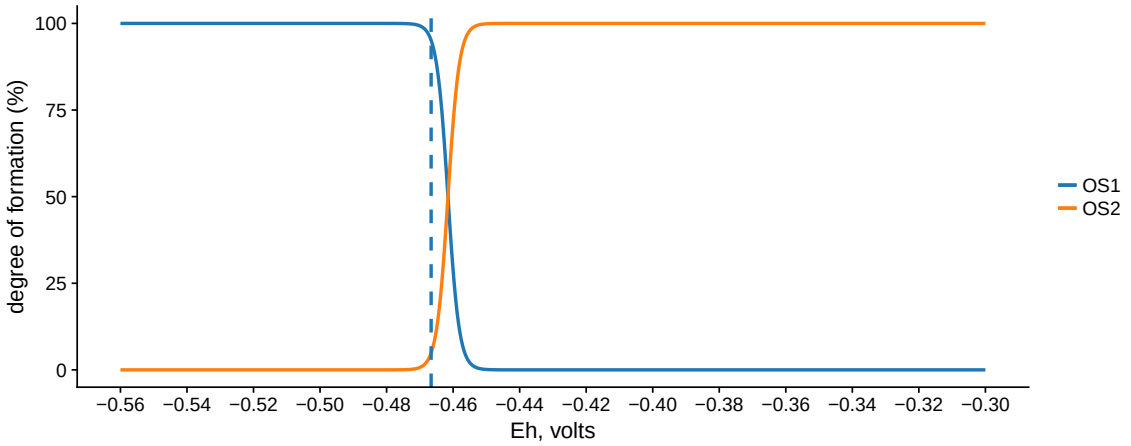


(c) EP4

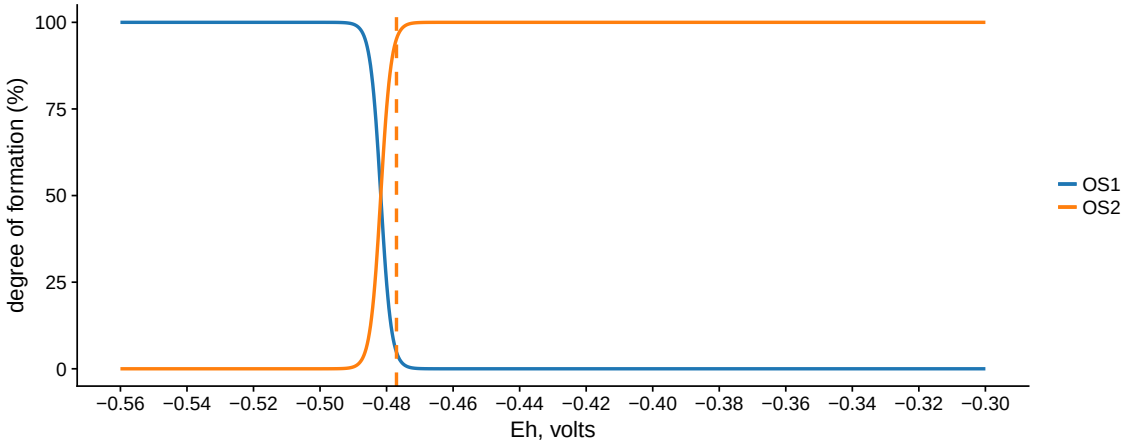


(d) EP5

Figure B.2: Metastable equilibrium abundance of sample-averaged free alkyl chains of Empress Pool predicted across an Eh gradient at the temperatures and bicarbonate, ammonium, and proton concentrations measured at (a) EP1, (b) EP3, (c) EP4, and (d) EP5. The vertical dotted line indicates the alkyl chain-predicted $\mathrm{Eh}_{\mathrm{alkyl}}$ of the sample.

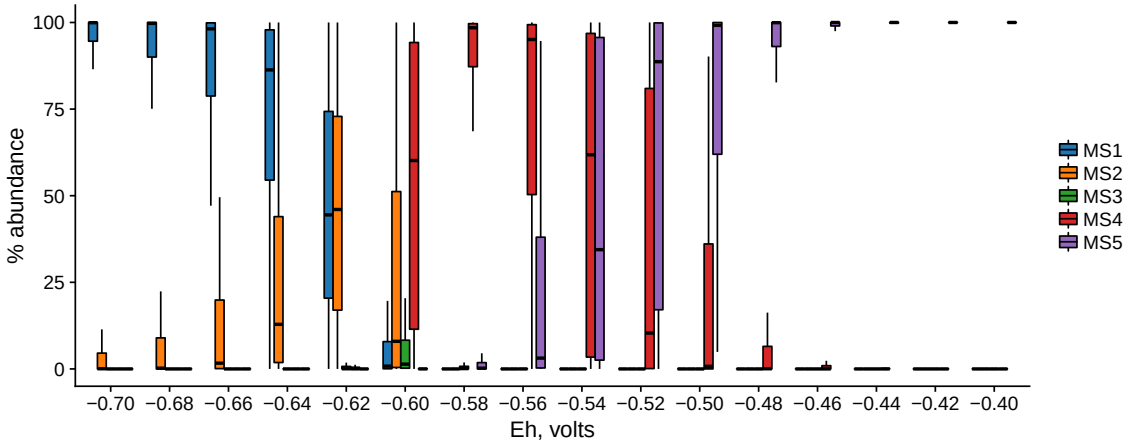


(a) OS1

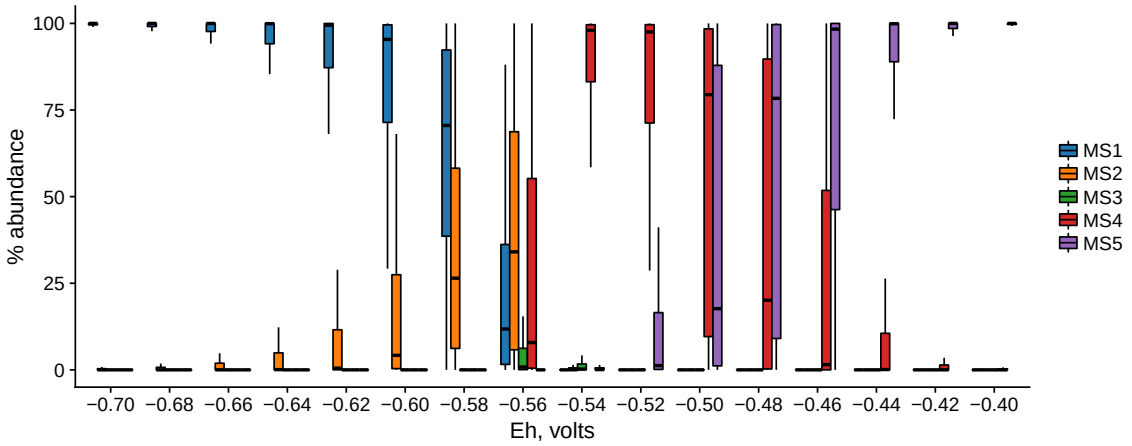


(b) OS2

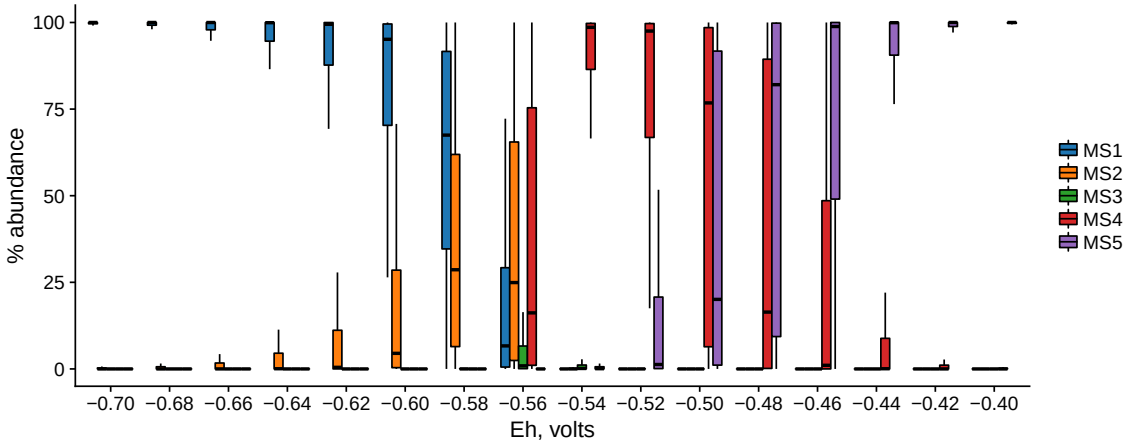
Figure B.3: Metastable equilibrium abundance of sample-averaged free alkyl chains of Octopus Spring predicted across an Eh gradient at the temperatures and bicarbonate, ammonium, and proton concentrations measured at (a) OS1 and (b) OS2. The vertical dotted line indicates the alkyl chain-predicted $\mathrm{Eh}_{\mathrm{alkyl}}$ of the sample.


100
75
50
25
0
% abundance
−0.70 −0.68 −0.66 −0.64 −0.62 −0.60 −0.58 −0.56 −0.54 −0.52 −0.50 −0.48 −0.46 −0.44 −0.42 −0.40
Eh, volts
MS1
MS2
MS3
MS4
MS5

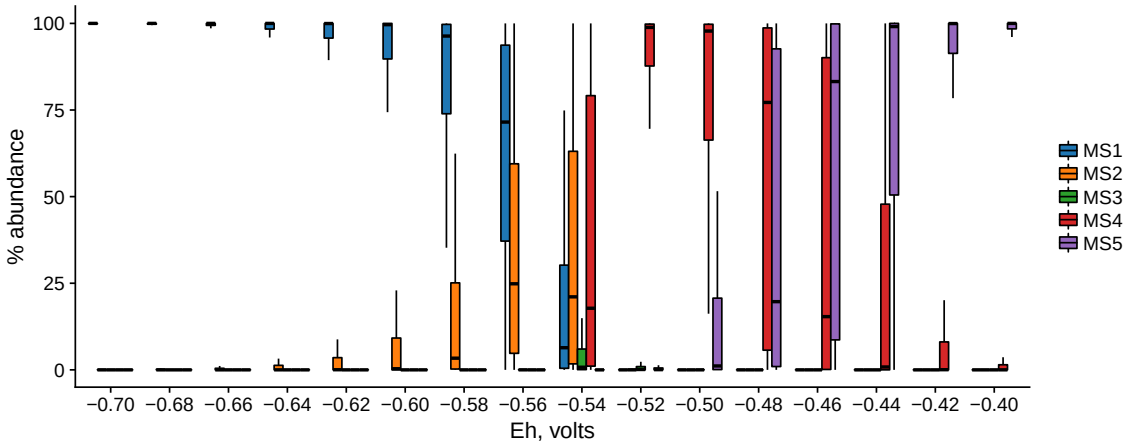
(a) MS1


100
75
50
25
0
% abundance
−0.70 −0.68 −0.66 −0.64 −0.62 −0.60 −0.58 −0.56 −0.54 −0.52 −0.50 −0.48 −0.46 −0.44 −0.42 −0.40
Eh, volts
MS1
MS2
MS3
MS4
MS5

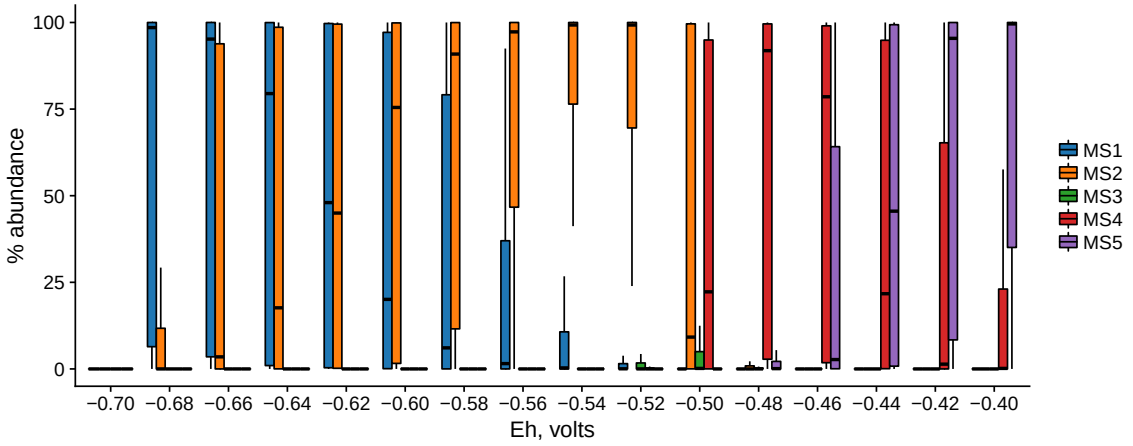
(b) MS2


100
75
50
25
0
% abundance
−0.70 −0.68 −0.66 −0.64 −0.62 −0.60 −0.58 −0.56 −0.54 −0.52 −0.50 −0.48 −0.46 −0.44 −0.42 −0.40
Eh, volts
MS1
MS2
MS3
MS4
MS5

(c) MS3

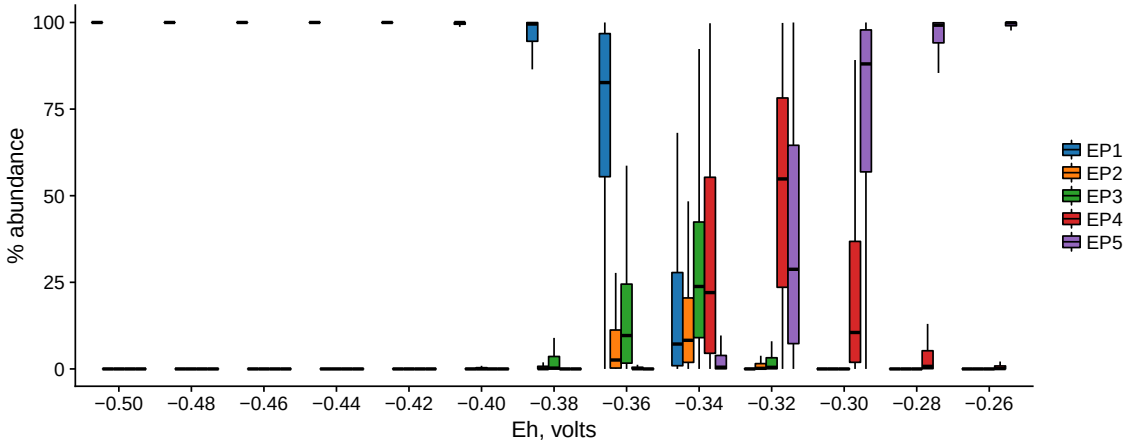

100
75
50
25
0
% abundance
−0.70 −0.68 −0.66 −0.64 −0.62 −0.60 −0.58 −0.56 −0.54 −0.52 −0.50 −0.48 −0.46 −0.44 −0.42 −0.40
Eh, volts
MS1
MS2
MS3
MS4
MS5

(d) MS4

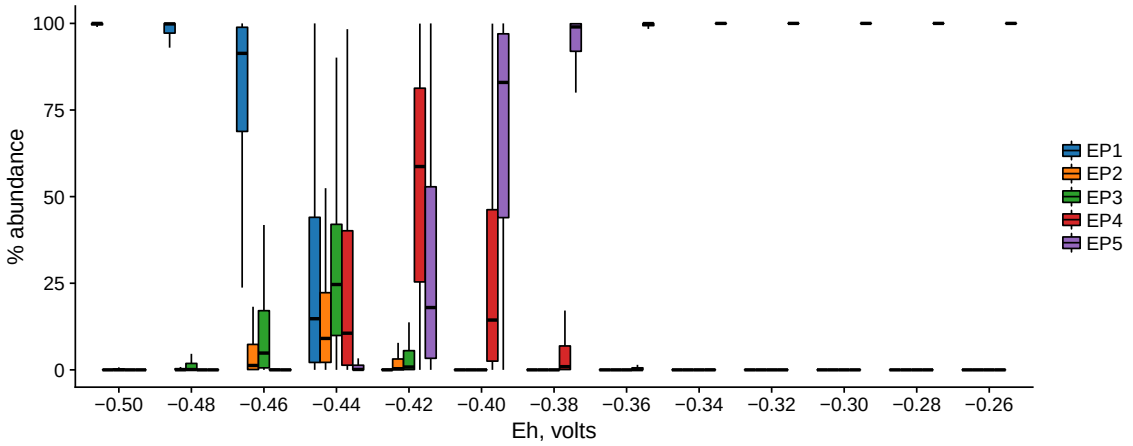


(e) MS5

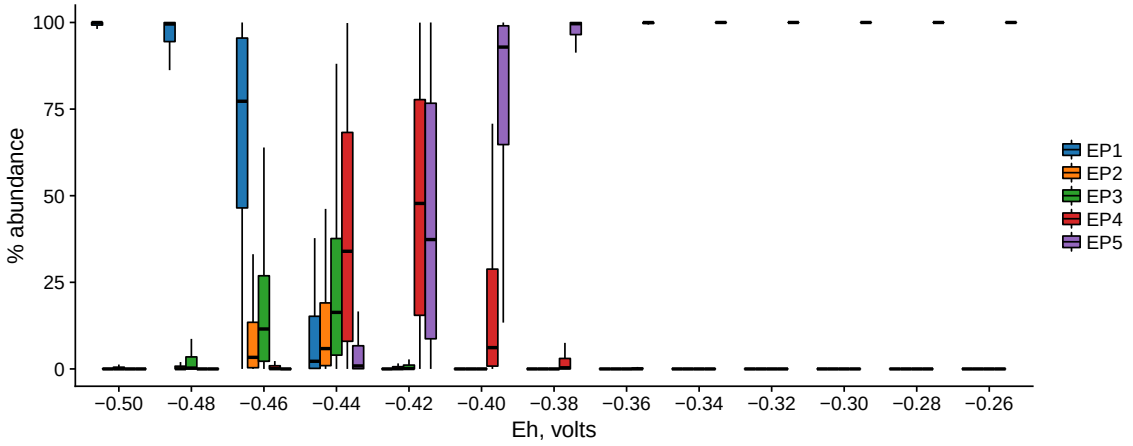
Figure B.4: Sensitivity analysis sample-averaged free alkyl chain metastable equilibrium abundance in Mound Spring samples (a) MS1, (b) MS2, (c) MS3, (d) MS4, and (e) MS5.


100
75
50
25
0
% abundance
−0.50
−0.48
−0.46
−0.44
−0.42
−0.40
−0.38
−0.36
−0.34
−0.32
−0.30
−0.28
−0.26
Eh, volts
EP1
EP2
EP3
EP4
EP5

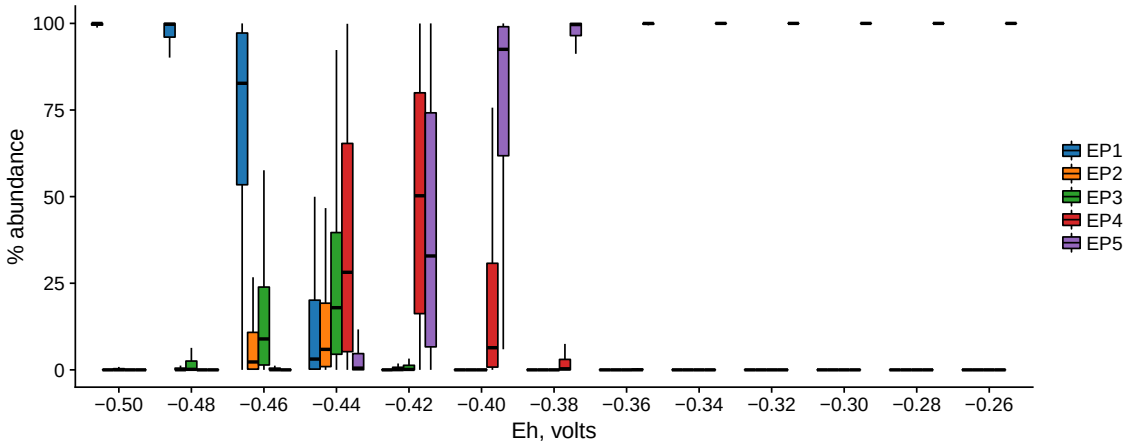
(a) EP1


100
75
50
25
0
% abundance
−0.50
−0.48
−0.46
−0.44
−0.42
−0.40
−0.38
−0.36
−0.34
−0.32
−0.30
−0.28
−0.26
Eh, volts
EP1
EP2
EP3
EP4
EP5

(b) EP3

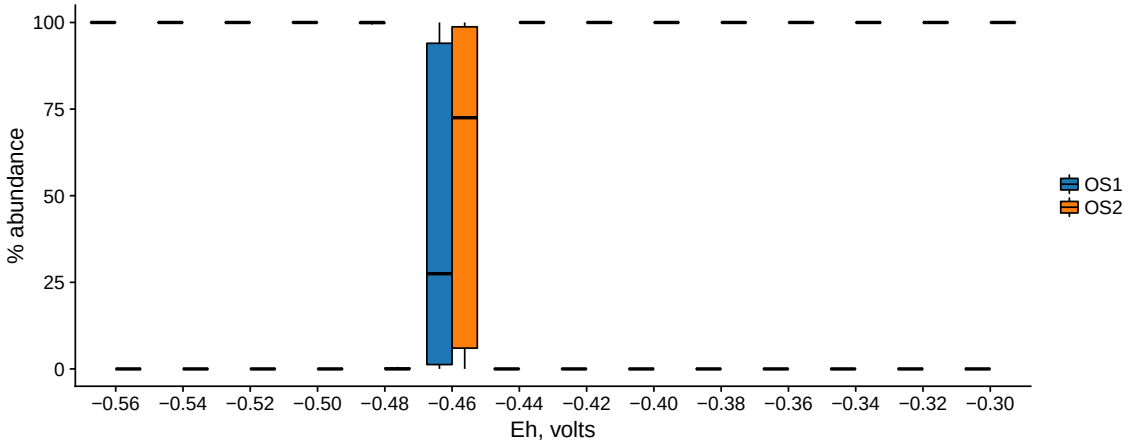


(c) EP4

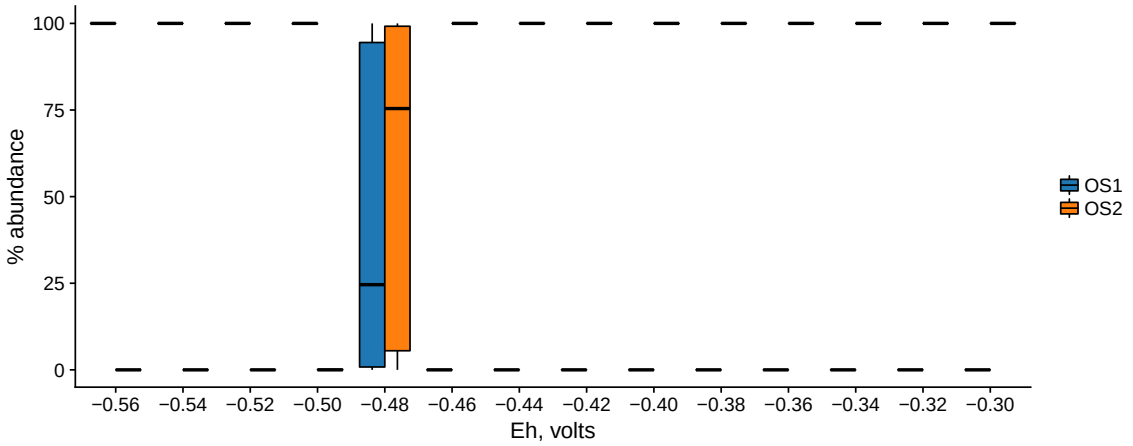


(d) EP5

Figure B.5: Sensitivity analysis sample-averaged free alkyl chain metastable equilibrium abundance in Empress Pool samples (a) EP1, (b) EP3, (c) EP4, and (d) EP5.



(a) OS1



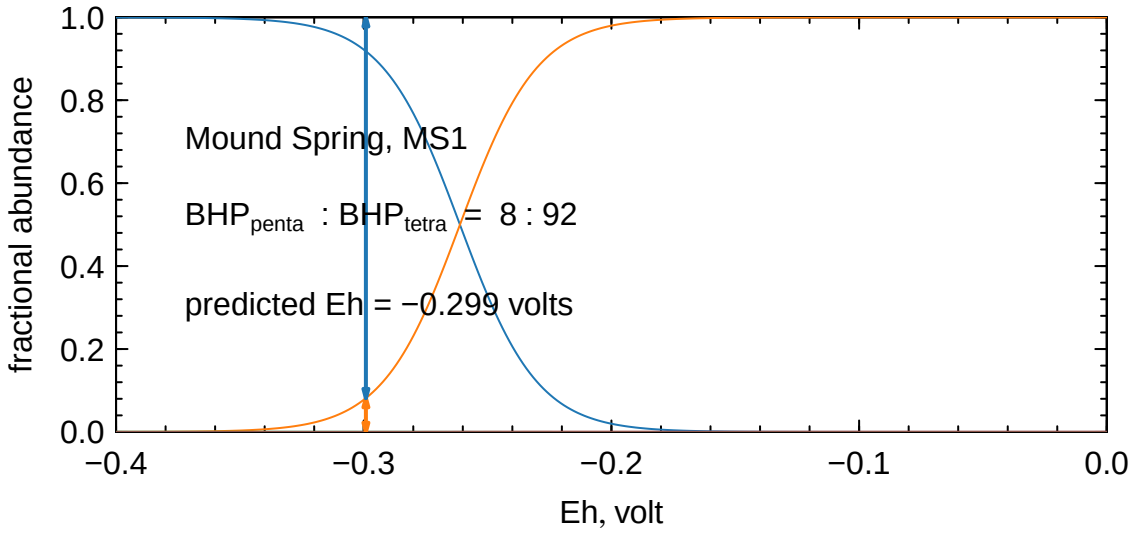
(b) OS2

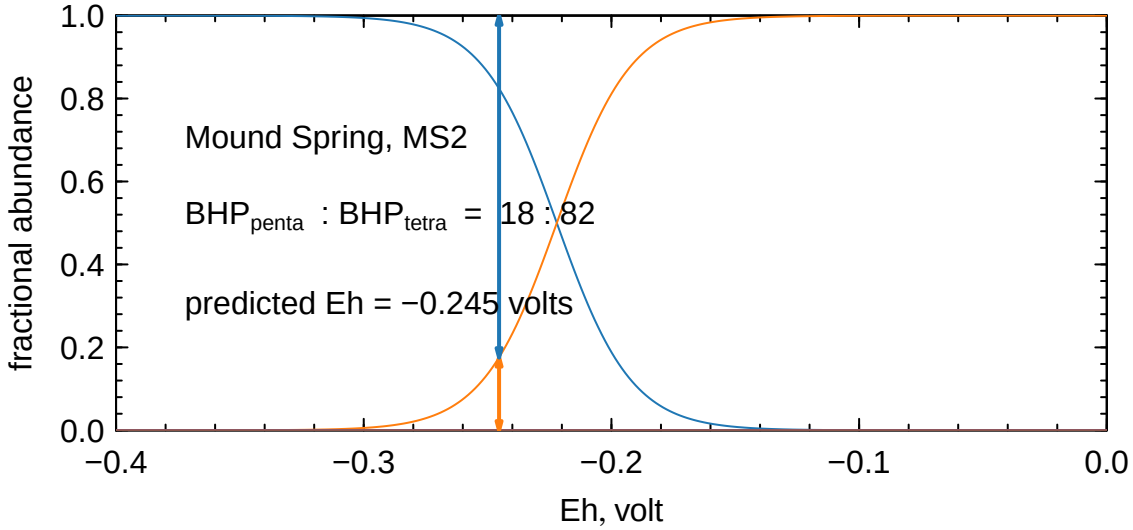
Figure B.6: Sensitivity analysis of sample-averaged free alkyl chain metastable equilibrium abundance in Octopus Spring samples (a) OS1, and (b) OS2.

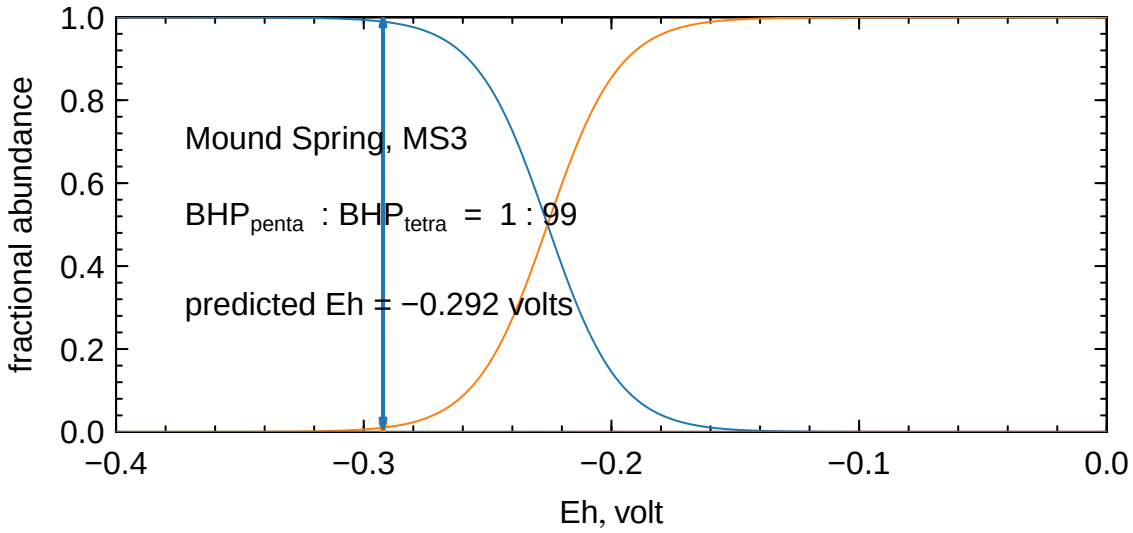
APPENDIX C

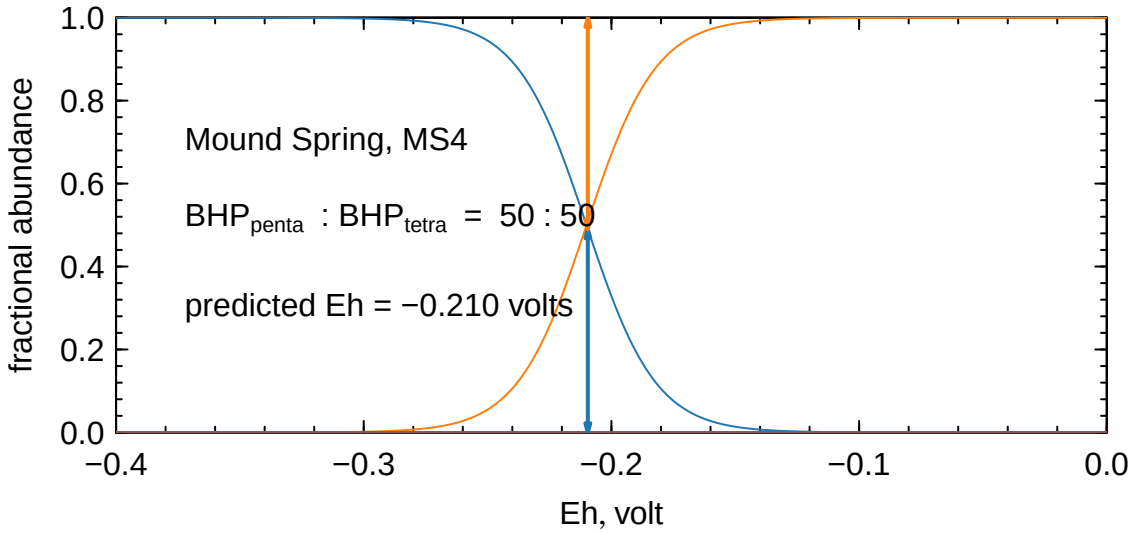
SUPPLEMENTARY FIGURES: PREDICTED EQUILIBRIUM ABUNDANCES OF BACTERIOHOPANEPOLYOL POLAR HEADGROUPS

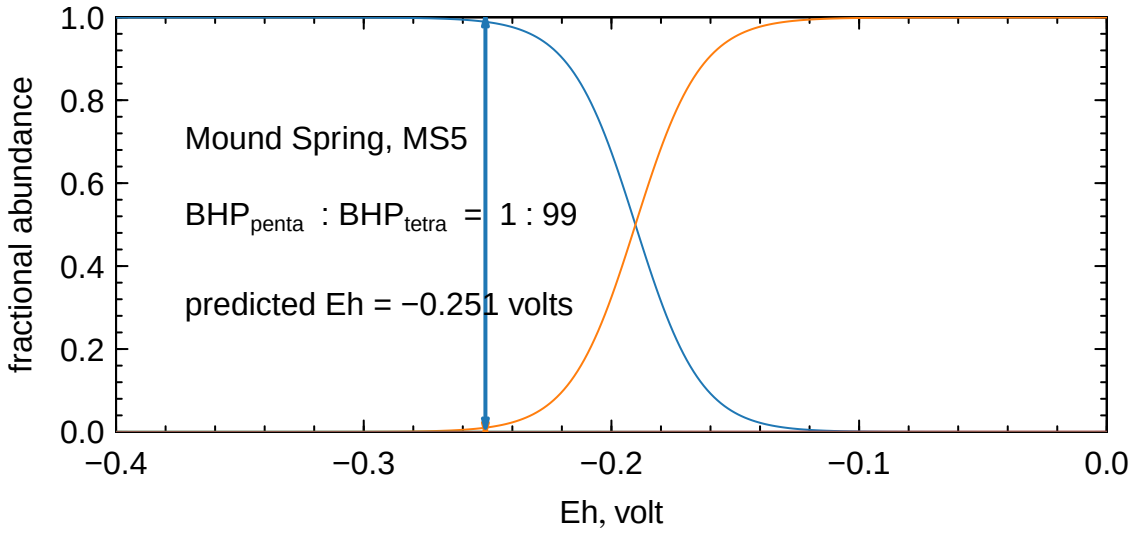
Equilibrium abundances. Predicted equilibrium abundances of tetrafunctional and pentafunctional BHP headgroups (BHP_{tetra} and BHP_{penta}, respectively) as a function of Eh are shown for Mound Spring, Empress Pool, and Octopus Spring in Figure C.1. Each subfigure shows a calculation at the temperature and basis species concentrations of a sample collected along the outflow channel of a spring. Figures for Bison Pool samples can be found in Chapter 4.

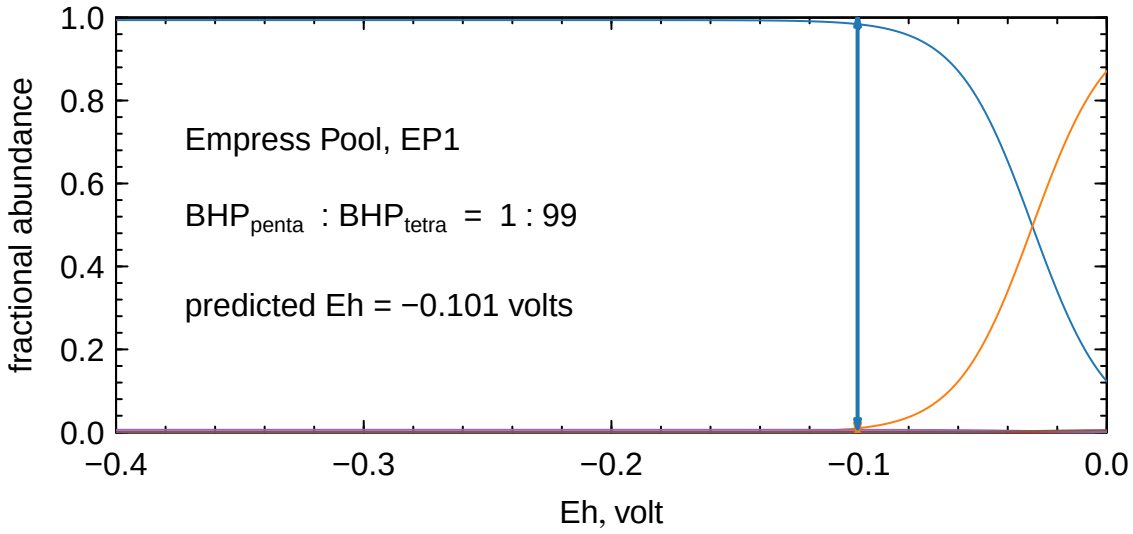

1.0
0.8
0.6
0.4
0.2
0.0
fractional abundance
Mound Spring, MS1
BHP$_{penta}$: BHP$_{tetra}$ = 8 : 92
predicted Eh = −0.299 volts
−0.4
−0.3
−0.2
−0.1
0.0
Eh, volt

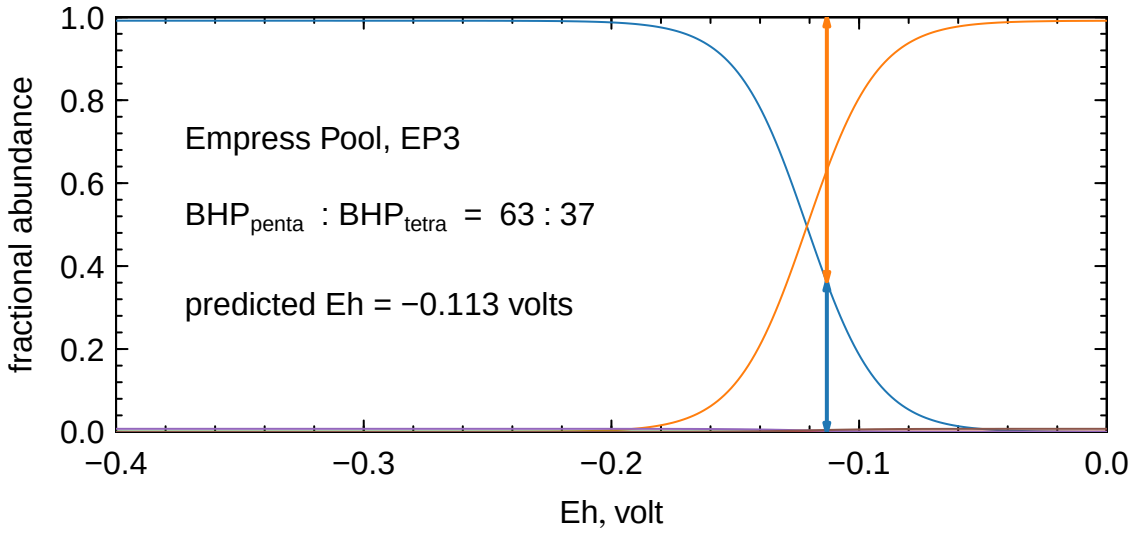

1.0
0.8
0.6
0.4
0.2
0.0
fractional abundance
Mound Spring, MS2
BHP$_{penta}$: BHP$_{tetra}$ = 18 : 82
predicted Eh = −0.245 volts
−0.4
−0.3
−0.2
−0.1
0.0
Eh, volt

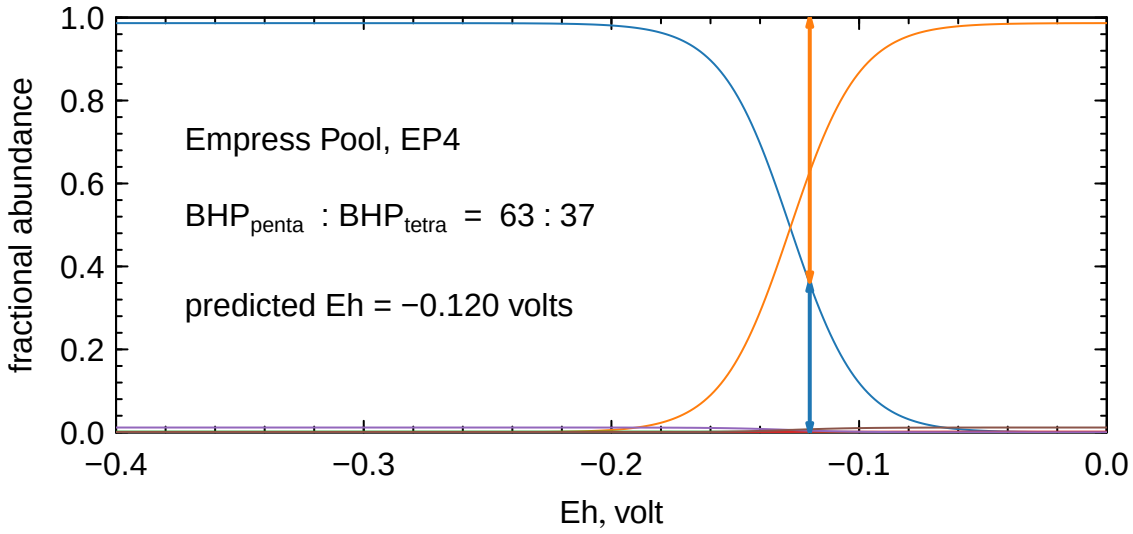

1.0
0.8
0.6
0.4
0.2
0.0
fractional abundance
Mound Spring, MS3
BHP$_{penta}$: BHP$_{tetra}$ = 1 : 99
predicted Eh = −0.292 volts
−0.4
−0.3
−0.2
−0.1
0.0
Eh, volt

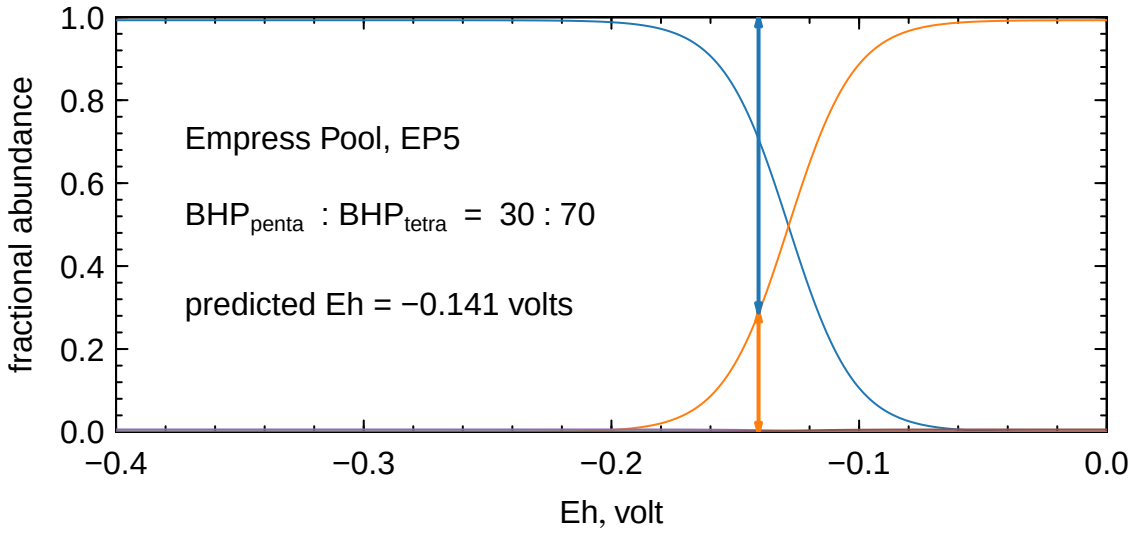

1.0
0.8
0.6
0.4
0.2
0.0
fractional abundance
Mound Spring, MS4
BHP$_{penta}$: BHP$_{tetra}$ = 50 : 50
predicted Eh = −0.210 volts
−0.4
−0.3
−0.2
−0.1
0.0
Eh, volt

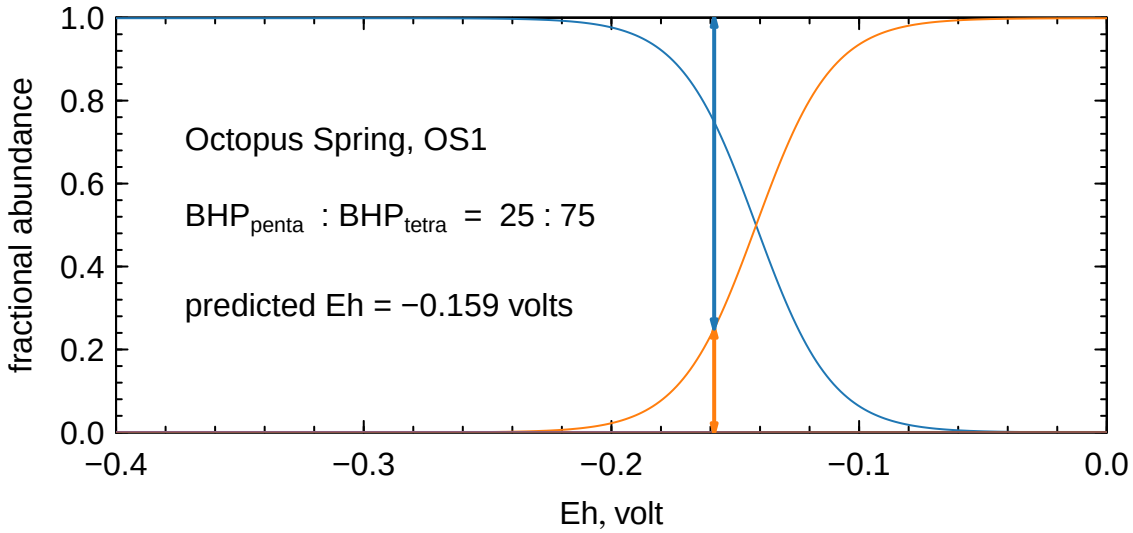

1.0
0.8
0.6
0.4
0.2
0.0
fractional abundance
Mound Spring, MS5
BHP$_{penta}$: BHP$_{tetra}$ = 1 : 99
predicted Eh = −0.251 volts
−0.4
−0.3
−0.2
−0.1
0.0
Eh, volt


1.0
0.8
0.6
0.4
0.2
0.0
fractional abundance
Empress Pool, EP1
BHP$_{penta}$: BHP$_{tetra}$ = 1 : 99
predicted Eh = −0.101 volts
−0.4
−0.3
−0.2
−0.1
0.0
Eh, volt


1.0
0.8
0.6
0.4
0.2
0.0
fractional abundance
Empress Pool, EP3
BHP$_{penta}$: BHP$_{tetra}$ = 63 : 37
predicted Eh = −0.113 volts
−0.4
−0.3
−0.2
−0.1
0.0
Eh, volt


1.0
0.8
0.6
0.4
0.2
0.0
fractional abundance
Empress Pool, EP4
BHP$_{penta}$: BHP$_{tetra}$ = 63 : 37
predicted Eh = −0.120 volts
−0.4
−0.3
−0.2
−0.1
0.0
Eh, volt


1.0
0.8
0.6
0.4
0.2
0.0
fractional abundance
Empress Pool, EP5
BHP$_{penta}$: BHP$_{tetra}$ = 30 : 70
predicted Eh = −0.141 volts
−0.4
−0.3
−0.2
−0.1
0.0
Eh, volt


1.0
0.8
0.6
0.4
0.2
0.0
fractional abundance
Octopus Spring, OS1
BHP$_{penta}$: BHP$_{tetra}$ = 25 : 75
predicted Eh = −0.159 volts
−0.4
−0.3
−0.2
−0.1
0.0
Eh, volt

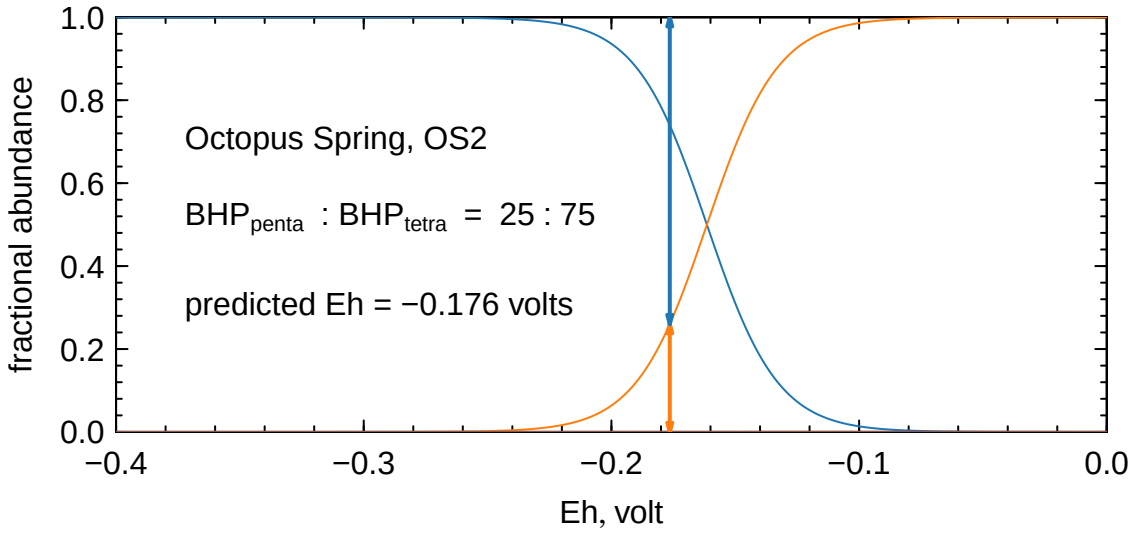


Figure C.1: Predicted metastable equilibrium abundance of tetrafunctional (blue) and pentafunctional (orange) BHP headgroups in Mound Spring, Empress Pool, and Octopus Spring samples compared to observed ratios (blue to orange vertical arrow lengths) with corresponding BHP headgroup-predicted Eh indicated for each.

APPENDIX D

THERMODYNAMIC EQUATIONS OF STATE

$$
\begin{aligned}
\Delta G^{\circ}_{aq} =& \Delta_f G^{\circ}_{aq} - S^{\circ}_{aq}(T - T_r) - c_1\Big[T \cdot \ln\frac{T}{T_r} - T + T_r\Big] \\
& - c_2\Big\{\Big[\Big(\frac{1}{T-\theta}\Big) - \Big(\frac{1}{T_r-\theta}\Big)\Big]\Big(\frac{\theta - T}{\theta}\Big) - \frac{T}{\theta^2}\ln\Big(\frac{T_r(T-\theta)}{T(T_r-\theta)}\Big)\Big\} \\
& + a_1(P - P_r) + a_2 \ln\Big(\frac{\Psi + P}{\Psi + P_r}\Big) + \Big(\frac{1}{T-\theta}\Big)\Big[a_3(P - P_r) + a_4 \ln\Big(\frac{\Psi + P}{\Psi + P_r}\Big)\Big] \\
& + \omega\Big(\frac{1}{\epsilon} - 1\Big)
\end{aligned}
\tag{D.1}
$$

The revised Helgeson-Kirkham-Flowers (HKF) equation of state used in this work was adapted from Equation B25 in Shock *et al.* (1992). Solvation terms not expected to have a significant effect on results within the range of pressure and temperature considered in this study (1 bar and < 100°C) were omitted. ΔG°_{aq} indicates the apparent standard partial molal Gibbs free energy of formation of the aqueous species from the elements at the temperature and pressure of interest. S°_{aq} and $\Delta_f G^{\circ}_{aq}$ stand for the standard partial molal entropy of the species and Gibbs free energy of formation of the aqueous species from the elements at 298.15K and 1 bar, respectively. a_1, a_2, a_3, a_4, c_1, c_2, and ω designate HKF equation of state parameters that are species-dependent. T and P stand for the temperature and pressure of interest. T_r and P_r represent the reference temperature and pressure of 298.15 K and 1 bar. θ and Ψ designate solvent parameters equal to 228 K and 2600 bar, respectively. ϵ stands for the dielectric constant of water at the temperature and pressure of interest.